

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2005, 20:29:58 ; Search time 51 Seconds
(without alignments)
1961.521 Million cell updates/sec

Title: US-10-677-980-2

Perfect score: 6481

Sequence: 1 MKGFNYFLIPLFLYNYI.....VOETNISDYSEYNYKRMV 1210

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.*

1: /cgm2_6/ptodata/1/iaa/5_COMB.pep.*

2: /cgm2_6/ptodata/1/iaa/6_COMB.pep.*

3: /cgm2_6/ptodata/1/iaa/H_COMB.pep.*

4: /cgm2_6/ptodata/1/iaa/PCUTS_COMB.pep.*

5: /cgm2_6/ptodata/1/iaa/RE_COMB.pep.*

6: /cgm2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1557	24.0	1435	1	US-08-568-459A-4
2	1557	24.0	1435	1	US-08-487-826B-4
3	1557	24.0	1435	2	US-09-210-288-4
4	1557	24.0	1435	2	US-10-153-273-4
5	768	11.9	1115	6	5198347-6
6	761.5	11.7	749	1	US-08-568-459A-6
7	761.5	11.7	749	1	US-08-487-826B-6
8	761.5	11.7	749	2	US-09-210-288-6
9	761.5	11.7	749	2	US-10-153-273-6
10	754	11.6	1115	1	US-08-568-459A-2
11	754	11.6	1115	1	US-08-487-826B-2
12	754	11.6	1115	2	US-09-210-288-2
13	754	11.6	1115	2	US-10-153-273-2
14	655	10.1	778	6	5198347-4
15	601.5	9.3	3542	2	US-10-087-013-2
16	498.5	7.7	3060	1	US-08-487-826B-14
17	495.5	7.6	2710	1	US-08-568-459A-12
18	495.5	7.6	2710	1	US-08-487-826B-12
19	495.5	7.6	2710	2	US-09-210-288-12
20	495.5	7.6	2710	2	US-10-153-273-12
21	432	6.7	1507	2	US-08-929-329-5
22	363	5.6	407	2	US-10-087-013-8
23	360.5	5.6	921	1	US-08-568-459A-8
24	360.5	5.6	921	1	US-08-487-826B-8
25	360.5	5.6	921	2	US-09-210-288-8
26	360.5	5.6	921	2	US-10-153-273-8
27	353.5	5.5	2182	1	US-08-487-826B-16

Sequence 7, Appli
Sequence 9, Appli
Sequence 11, Appli
Sequence 10, Appli
Sequence 10, Appli
Sequence 10, Appli
Sequence 10, Appli
Sequence 10, Appli
Sequence 10, Appli
Sequence 10, Appli
Sequence 6, Appli
Sequence 15, Appli
Sequence 27, Appli
Sequence 15, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 16, Appli
Sequence 28, Appli

ALIGNMENTS

RESULT 1

US-08-568-459A-4
; Sequence 4, Application US/08568459A
; Patent No. 5849306
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellem, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,459A
; FILING DATE: 07-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-0176
; TELEFAX: (619) 235-8550
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1435 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
; US-08-568-459A-4

Query Match 24.0%; Score 1557; DB 1; Length 1435;

	US-08-487-826B-4	ORGANISM: Plasmodium falciparum	
	Query Match 24.0%; Score 1557; DB 1; Length 1435;		
	Best Local Similarity 28.5%; Pred. No. 1.2e-98;		
	Matches 410; Conservative 214; Mismatches 467; Indels 348; Gaps 42;		
QY	1 MKGYFNIYFLIPLIFLYNIRINESIGRTLYNRODESSDLSRVNSPELNNHNTNYDS 60		
DB	1 MKCNSIYFPASFFVLYFAKARNBYDI-----KENEFLDVYKEGFNELOKKYGNVQKT 55		
QY	61 DYEDVNKLINGFVENKSVKCKRSISPINN---KTKSY---DIIPPSYRNDKENS L 112		
DB	56 D-----KKIFTIEKN-----LDILNSKFNKWKS YGTPDNIDKMSLIN----- 96		
QY	113 SENEONGTNSNFANTSEISIGNDKNQ-----YTFIQGRTH----- 150		
DB	97 KHNBEFMFNYYQSFLSTS--SLIKQKYVPINAVRVSRIILSFLDRINNGBNTSSNNEV 154		
QY	151 LFACGIKRKSIKIWI CRENSEKIT-VCVPRKIQLCVANFLNSRLTMEKFKPI LISVNT 209		
DB	155 LSNCREKRGMEKWDCKKGNNYVICIPDRITQLCIVNLISIITYKTEWKDHPIEASKI 214		
QY	210 EAKLLYNKNEGDPISFCNELRNSPFSRFIGDDMDPGGTVRVGYNITKESDYKE 269		
DB	215 ESOLLKXNDKNYKSCFNDLKNSFLDYLGLAMGDMDFGYSYKAENKI QEVFKGAHE 274		
QY	270 KNVEKLANNIKCEWEKKNKANLNWHMIVNHKGISKECALIIPABEPQINLMIKEWENFLM 329		
DB	275 ISEHKLIKPKRWEMFEKLEWAMLSEHKNNIN-NCKNIPOBELQITQWIKEHGFEFL 333		
QY	330 EKRRFLNLIKOVENKKYKAEFGCGRLPCSSYTSFMKKSKTQMEVLTNLYKKNGSGVDK 389		
DB	334 ERDRSKLPKSKCNNTLYEAKEKICIDCMKYRDWIIIRSKFEWHTLSKEYETQK--VPK 391		
QY	390 NNFLNDLFK-KNNKD--LDDEPKN-EKEYDDLDCRYTATIIKSFPLNGPAKNDVDIASQ 445		
DB	392 ENAENYLIKISEKNDAKVSLLLNNCAEYSKYCDKHRTTLTVKSVLNGNDNTIKEKREH 451		
QY	446 INVNDLRGFGCNYS--NNEKSWNCCTGFTNKPFGTCPPRRQTCLGRVTLLHRGHEED 503		
DB	452 IDLDPFSGCDKNSVDNTNKKWCNKNPYILLTDKDVCPRRQBELCLGNTDRIYDKMLLM 511		
QY	504 KYEHLGASIYBAQLLYKYRKEDNALCSIIIONSADLADIIGSDI IKDYTGKMEEN 563		
DB	512 IKEXHIAIAYESRIKRYKNDKEVCIINTKPTADIRDIIGDTYWNDSLNRKLGVK 571		
QY	564 LNKVNDKGRNESIKIFREKWWDENKENVKWSAVLNKETCEDYKQFKQIPOPURWP 623		
DB	572 INTNSKYHRNKNNDKLFRDENWVIKDVNVVVISWVFQDKTVCKE-DDIENTIPQQPFWP 630		
QY	624 KEWGDDFCBKREKILYSFESFKVCCKKDCENTCKNSCYKWKWIDLKSEYEKVQV--- 680		
DB	631 SEWGDDYCQDKYMI----ETLVCKECPKCPDDNCKNSCYKWKWISKKEEYNKQAKY 687		
QY	681 DKYTQDKNKMVDNIDEVKNKANYLLKESKCKVDNFDPKIENESPNEYDMCKKCDE 740		
DB	688 QEYQKGNMYMI---SEPKSIKPEVYLKYSKESKSNLNFEDFPEKEELHSYKKNKTCWCP 744		
QY	741 IKYL-----NK----- 746		
DB	745 VKDVPISIRNNEOTSQBAVPEENTEIAHRTETPISBGPKGNEKERDDDSLSKISVSP 804		
QY	747 -----IKYKPT----- 752		
DB	805 ENSRPETDAKOTSNLLKLGVDVIMSPKAVITGSSPNNDINVTQEGDNISGVNSKPLSDVV 864		
QY	753 ----KHDIYDID-----TFSDTFGDTGPIS----- 774		
DB	865 RPDKKELEDQNSEETVWNHISKPSINNGDDSGSATVSESSSNTGLSIDDDRNG 924		
QY	775 -----NANINEQGSGKDTNNTGNETSDSP---VSHEPESDAAINV 812		

Db	745	VKDVPISIRNNQTSQEA	VPENTEIAHRTETP	TSISGPGKNGEQRDD	SLSKISVSP	804
Qy	747	-----IKPKT-----	-----	-----	-----	752
Db	805	ENSRPETDAKDTSNLLK	KGVDVISMFKAVIGSS	PNDNINVT	EQDNIISGVNSKPLSDDV	864
Qy	753	---KHDIYDID---	-----	-----	-----	774
Db	865	RPDKLELDQNSDESEET	VVNHISKSPSINNGDD	SGSGSATVSESSSNT	GLSISDDDRNG	924
Qy	775	-----	-----	-----	-----	812
Db	925	DTFVRTQDTANTEDVIR	KENADKDEKGADEHST	SESLSPEEKMLTDNEG	CSLNH	984
Qy	813	EKL-----SGDESSE	TRGILDIN-----	-----	-----	846
Db	985	EEVKEHTSNDNVQSGGI	VMMVVEKELDTLENP	SSSLDEKGAHELS	SEPNLSSDQDMS	1044
Qy	847	NTQGSVNTSDITN---	GHSESLNR-----	-----	-----	892
Db	1045	NTFGPLDNTSEETTER	ISINNEYKVNREDERT	LTKEYEDIVLKSHMNR	RESDDGDELYDENS	1104
Qy	893	SHSDDSGSLTI--G	QVPSDENTQNTY-----	-----	-----	934
Db	1105	DLSTVNDSEDAEAKM	GNDTSEMSHNSQHI	ESDQQRNDMTVGD	LGTTHVQNEISVPV	1164
Qy	935	---SDDKINE-----	-----	-----	-----	967
Db	1165	TGEIDEKLRSESKIH	KABEERLSHTDIHK	INPEDRNSNTHL	KDIRNEENERHLTNQ	1224
Qy	968	TNIVSERRVNSHP	IRNGMANNHAHQYITQ	-----	-----	1021
Db	1225	INISQERDLQKGF---	HTMNNLHG	DGVSESRQINSHSH	EGNRQDRGSGNVLNMF	1280
Qy	1022	-----PKRNPSS	ENDHKNIQVNSRDT	KRVREETIKLSKON	CKNNYESMEYCTYSDE	1075
Db	1281	NFNFNTPSRVNL---	YDKLUDL	OLYERNDSNTHK	ELKLABINCKENELSV	1337
Qy	1076	RNSPGPCSRERKGLCC	QISDYCLKYFNFY	SYEYINCYKSEIK	SPKYCKFKSGQSSI	1134
Db	1338	EEIPLKCTKTEKTRNL	CCAVDYCMSTYTYD	SEBYNCTKREP	DDPSYTCFRKEAPSSM	1396
RESULT 5						
5198347-6						
; Patent No. 5198347						
; APPLICANT: MILLER, LOUIS H.; ADAMS, JOHN H.; KASLOW,						
; DAVID C.; FANG, XIANGDOUG						
; TITLE OF INVENTION: DNA ENCODING PLASMODIUM VIVAX AND						
; PLASMODIUM KNOWLESI DUFFY RECEPTOR						
; NUMBER OF SEQUENCES: 27						
; CURRENT APPLICATION DATA:						
; APPLICATION NUMBER: US/08/554,837						
; FILING DATE: 20-JUL-1990						
; SEQ ID NO:6:						
; LENGTH: 1115						
5198347-6						
Query Match 11.9%; Score 768; DB 6; Length 1115;						
Best Local Similarity 24.5%; Pred. No. 2.7e-44;						
Matches 296; Conservative 178; Mismatches 416; Indels 316; Gaps 56;						
Qy	199	FKBIFILSVNTEAKL	VNKNKEGDKPSIF	CFNCLNLSFSDR	SSFIGDDMDFGNTDR	258
Db	22	YKDDFSIT-----	LINHEGKKYLILKR	-----	-----	59
Qy	259	IN--TKFSDYYKEK	VNKLNNIKKEWBE	--KNKANLWNHMI	VNHKGNISKECALIP	316
Db	60	LHFSQVNNVLLERT	ITETLLBCKEYVKG	NGYKYL-----	AKGH-----	105
Qy	317	NLMIKENENFLMEK	KRLFLNIDKDCV	ENKCYEACFGC	BRLPCS--SYTSPMK	375
Db	106	ERWLQGTNR-----	-----	-----	-----	149

Qy	376	L-----TNLYKKNKSGVQKQNFNLDFKQKQNKQDLDDFFPKNEKEYDDLCDCRY-----423
Db	150	IYGAHNFNGSMEYBGKGGDKTG-----EERKDGHEKTDKTDNGKGANNILVMDIYETSSN204
Qy	424	-----TATIISKFLNGPAKN-----DVD-----TASQINLVNDLRGF-----G455
Db	205	GOPAGTLDNVLEFVTGHEGHSRKNSNGGNPYDIDHKHTISSAT-IN-HAPLQNTWKN261
Qy	456	CNYK-SNNEKSWNCTGTFTNKFPGTCEBPPRQTLCLGRTYLLHHRGHEEDY-----KEH507
Db	262	CNYKRRERDWC-----NTKQDVCIPDRRYQLCMKELTNLVNNTDTNFRHDITPRKLY316
Qy	508	LLGASIYEA-----QLLKYYKKEDEVALCSIIQNSYADLADIIGKSGDIKYGKK559
Db	317	LKRKLIYDAAVEGDLKJANNRY-NKD-----FCKDIKWSLGDGFDIIMGDTMEBIGYGSV372
Qy	560	MEENLNV-NKDKKRNEBSLKIPEKWWDENKENVWMSAVLKNK-----ETCKDYDK612
Db	373	VENNLRSIFGTDEKAQR-----RKQWNESKAIWTAMYSVKRLLKGNFIWICKLVNA427
Qy	613	FQKIPQFLRPFKEWGDGPEKREKIKIYSPFSFKVEC-----KKQCDNBNTCKNKGCSEY665
Db	428	VNIEPQIYRITREWRGRDVSLEPTEV-----QKLKEKCDGKINYTDKVKCVPPCQNAKSY484
Qy	666	KWITDLKSEYKQVDKYTKDKNKMYDNIDEVKNKEANV-----YLKESKCKOVNFD720
Db	485	QOWITRKQNDVLNSKFI SVKNAB-----KVQTAGIVTPYDIIQLBELDFNEVAFE536
Qy	721	DKIPNESPYEDMCKRDEBIKYLNEIKYKPKRHDYDIDFTSDTFG-DGTPFI-----772
Db	537	NEI-NKRDGAYIELC-----VCSVEEAK-KNTQEVWTVNDNAAKSQATNSNPISQPDVSS589
Qy	773	-----SINANINEOOSKDTNTGNSETSD-----SPVSH-----PSSDAINV812
Db	590	KAEKVPBGDTHGNVN-----SQDSSSTTKAVTGDCQNGQNTPAESDVQVORDIAESVSAKV646
Qy	813	E-----KLSGD-----ESSSETRGILDINDPSVTNNVNEVHDASNTQGSV-----852
Db	647	DPQKSVKREDDTASVTGLAEAGENLGA NSRRSESTVEANSPCDDTVNSASIPVVSGE706
Qy	853	-----SNTSDITNGHSSSLNRTTNAQDIKIRSG-----NEQSDNQENSSHSS896
Db	707	NPLVTPYNGLRHSKDNSDSCPAESMANPDNSK-----GETGKGQDNMAKATKDSNSS762
Qy	897	DNSGSLN-----IQOVSEBNTQNTYDSQNPHRDTPNALASLPSDDKINIEIG944
Db	763	DGISSATGDTTDAVREINKVPEDRKTVGSKOGGEDNSANKDAATVVGEDIRINSA822
Qy	945	FDSSRD-----SENG-----RGDTTSN--THDVRRNTIV971
Db	823	GGSTNDRSKNDTEKNKGASTPDSKQSEDTALSKTESLESTESGDRTTNDTNSLENKNG882
Qy	972	SERVNSHDFTRNGWANNHAHVIT-----QIENNGIIRCOEBSAGNSVNYKDNPK1023
Db	883	KEKDLQKHDFAKSDTPEEFPNSDQTTDAEGHDRDSIKND---KAERKHKMNKOTFTKQTN939
Qy	1024	RSNFESSEN-----DHKKNIQEYNSRDTKRVREIKLSKQNKCNNEYSMEYCTYSDERNSSP1080
Db	940	SHLANSNNLSNGKLDIKEYKYRDVKATREDIILMSSVRKCNNNISILEYCNVSEDKISS-998
Qy	1081	GPCSRERKKLCCOISDYCLKYFNFYISIEYNCIKSEIK3PEYKCFKSGEQSSIPYPAAG1140
Db	999	NTCSREKSNLCCISDFCLNYFDVYSYELSCMKKEFEDPSYKCFTKGGFKDKTYPAAA1058
Qy	1141	GILVIVILLSSASRGMKSNEEYDIGESNTTEATFEE--NNYLKLSRI-----FNOEVOET1194
Db	1059	GALLIILLILLIAS-RKMLKNDS-----EATNFEEFYCDNIHRIPLMPNNIEHMQPS1109
Qy	1195	NISDYS1200
Db	1110	TPLDYS1115

RESULT 6
 US-08-568-459A-6
 ; Sequence 6, Application US/08568459A
 ; Patent No. 5849306
 ; GENERAL INFORMATION:
 ; APPLICANT: Sim, Kim L.
 ; APPLICANT: Chitnis, Chetan
 ; APPLICANT: Miller, Louis H.
 ; APPLICANT: Peterson, David S.
 ; APPLICANT: Su, Xin-zhaun
 ; APPLICANT: Wellemans, Thomas E.
 ; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
 ; AND PLASMODIUM FALCIPARUM ERYTHROCYTE
 ; NUMBER OF SEQUENCES: 37
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Knobbe Martens Olson & Bear
 ; STREET: 620 Newport Center Drive 16th Floor
 ; CITY: Newport Beach
 ; STATE: California
 ; COUNTRY: US
 ; ZIP: 92660
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/568,459A
 ; FILING DATE: 07-DEC-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Israelse, Ned
 ; REGISTRATION NUMBER: 29,655
 ; REFERENCE/DOCKET NUMBER: NIH121.001CPI
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (619) 235-8550
 ; TELEFAX: (619) 235-0176
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 749 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: Plasmodium falciparum
 ; US-08-568-459A-6

[illegible]

RESULT 7

US-08-487-826B-6
: Sequence 6, Application US/08487826B
: Patent No. 5993827
: GENERAL INFORMATION:
: APPLICANT: Sim, Kim L.
: APPLICANT: Chitnis, Chetan
: APPLICANT: Miller, Louis H.
: APPLICANT: Peterson, David S.
: APPLICANT: Su, Xin-zhaun
: APPLICANT: Welles, Thomas E.
: TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
: NUMBER OF SEQUENCES: 45
: CORRESPONDENCE ADDRESS:
: ADDRESSER: Knobbe Martens Olson & Bear
: STREET: 620 Newport Center Drive 16th Floor
: CITY: Newport Beach
: STATE: California
: COUNTRY: US
: ZIP: 92660
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/487,826B
: FILING DATE: 10-SEP-1993
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Israelsen, Ned
: REGISTRATION NUMBER: 29,655
: REFERENCE/DOCKET NUMBER: NIH121.001CPI
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 235-8550
: TELEFAX: (619) 235-0176
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:

Db 373 VENNLRISFGTDEKAQOR-----RKQWNNESKAQIWTAMMYSVKRLKGNFIWICKLNVA 427
Qy 613 FQKIPQFLRFKFWGDDFCBKRKEKIXSFSEFKVE-----KKKDCDENTCKNKCEY 665
Db 428 VNIETQIYRWIREGRDYVSELPTFV---QKLKEKCDGKINYTDKVKVPPQCNACKSY 484
Qy 666 KKWIDLKSEYKQVDKYTKDNKKNKMYNDIDEVKNKEANV-----YLKEKSKCKOVNFD 720
Db 485 DOMITRKNQWDLVSNKFIYSQNAE-----KVQTAGIWTVPYDILKQBLDFEVAFAE 536
Qy 721 KKLFPNEPVEDMCKKDEIKYLNEIKYPTKHDIYDIDTFSDTFG-DGTPI----- 772
Db 537 NEI-NKRDGAYIELC-----VCSVEEAK-KNTQEVVTNVNDAKSAQATNSNPISQPVDS 589
Qy 773 -----SINANINEQSGKDTNTGNSETSD-----SPVSHE-----PESDAAINV 812
Db 590 KAELKVPDSTHGNV---SQGDSSTTKAVTGDQNGNQTPAESDVQBSIDAESVSAKV 646
Qy 813 E-----KLSGD-----ESSSETRGILDIINDPSTVNNVNEVHDASNTQGV----- 852
Db 647 DPQKSVSKRSDDTASVTGIAEAGKENLGASNSRSESTVEANSPGDDTVNSASIPVVSGE 706
Qy 853 -----SNTSDITNGHSESSLNRTTNAQDIKIRSG-----NEQSDNQENSHSS 896
Db 707 NPLVTPYNGLRHSKNSDSGPAESMANPDSNK-----GETGKGQNDMAKATKDSNSS 762
Qy 897 DNGSLT-----ICQVPSEDNTQNTVDSQNPHRDTPNALASLPDDKINEIEG 944
Db 763 DGTSSATGTTDAVDREINKGVPEDRDKTVGSKDGGEDNSANKDAATVVGEDIRENSA 822
Qy 945 FDSGRD-----SENG-----RGDTTSN---THDVRRTNIV 971
Db 823 GGSTNDRSKNDTEKNGASTPDSKQSEDATASLTESLESTESGDRTTNDTNSLENKNG 882
Qy 972 SERGVNSHDFIRGMANNNAHQVIT-----QIENNGIIRQOBESAGNSVNYKDNPK 1023
Db 883 KENDLOQKDFKSNTPNEEPSNDGTTAEGHDRDSIKND---KAERKKHMKDFTTKNTN 939
Qy 1024 RSNPSEN---DHKKNTQEYNSRTRKVRBEIILSKQNKCNNEYSMEYCTYSDERNSSP 1080
Db 940 SHLANSNNLSNGKLDIKEKYRDVKATREDIILMSSVRKCNNNISLEYCNSVEDKLSS- 998
Qy 1081 GPCSRERBKLCQOISYCLKYFNFYIYNYCNKISEIKSPYKCFKSEG-QSSIPYFAA 1139
Db 999 NTCRSREKSKNLCCSISDFCLNYFDVSYEYLSQMKKEFEPSYKCFKGGFKIDKTYFAA 1058
Qy 1140 GGLIWLIVLLSSASRMKSNEEYDIGESNEATFEF-NNYLAKLSRI-----FNQEOVE 1193
Db 1059 AGAL--LILLIARRKMIKNDSE-----EATNFEEYCONIHRPIPLMPNNIEHQ 1108
Qy 1194 TNSDYS 1200
Db 1109 STPLDYS 1115

RESULT 11

US-08-487-826B-2
; Sequence 2, Application US/08487826B
; Patent No. 5993827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellem, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; NUMBER OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor

CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1115 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium vivax
US-08-487-826B-2

Query Match 11.6%; Score 754; DB 1; Length 1115;
Best Local Similarity 24.5%; Pred. No. 2.5e-43;
Matches 296; Conservative 177; Mismatches 416; Indels 318; Gaps 57;

Qy 199 FKEIFILSVNTEAKLYNNKNEKDPISFCNELRNSPDSFRSSFIGDDMDPGGNTDRVKGY 258
Db 22 YKDDFSIT-----LYNHEGKKYLIILKR-----KLEKANNRVDVNCFF 59
Qy 259 IN-TKESDYKKNVKNKLANIKKEWE-KNKANLWNNHMIWNHKNISKECALIPASEPQI 316
Db 60 LHFSQNNVLLERTIETLLECKNEYVKGNGYKL-----AKGH---HC-----VEEDNL 105
Qy 317 NLWIKENENFLMEKKRRLFLNINOKCVENKYEACFGGRLPCS-SYTSFMKSKTKQMEV 375
Db 106 ERWLQGTNER-----RSEENIKYK--YGVTELKIKYAQMNGKRSRILKES 149
Qy 376 L-----TNLYKKNSGVDKNFLNDLFGKNNKNDLDDFFKNEKEYDDLCDCRY----- 423
Db 150 IYGAHNFSGNSYMEGKDGDKTG-----BEKQGEHKTSDKTDNGKANNLVLMDYETSSN 204
Qy 424 -----TATIKSGFLNGPAKN-----DVD-----IASQINVDNLGRF-----G 455
Db 205 GQPAQGLDNLVLEPVTCHEGNSRKNSSNGNPGYDIDHKYISSAI-IN--HAPLQNTVMKN 261
Qy 456 CNYK-SNNEKSNMCTGTFTNKPFGTCPPRRQTLCLGRTYLLHRGHEEDY-----KEH 507
Db 262 CNYKRRERDWDG-----NTKKDVCIPDRRYQLCKMELTNLVNNTDTNPHRDTTPRKLY 316
Qy 508 LLGASIEA-----QLLKYYKEKEDENALCSIIQNSYADLADIIGSDIIDKYGGKK 559
Db 317 LKREKLYDAAVEGDDLKLNRY-NKD---FKDIRWSLGLDGDIIIMGTDMGIGYSKV 372
Qy 560 MEENLNK-NKDKRNEESLIKIFREKWDENKENVKMSAVLNK-----ETCKDYDK 612
Db 373 VENNLRSIFGTDEKAQOR-----RKQWNNESKAQIWTAMMYSVKRLKGNFIWICKLNVA 427
Qy 613 FQKIPQFLRFKFWGDDFCBKRKEKIXSFSEFKVE-----KKKDCDENTCKNKCEY 665
Db 428 VNIETQIYRWIREGRDYVSELPTFV---QKLKEKCDGKINYTDKVKVPPQCNACKSY 484
Qy 666 KKWIDLKSEYKQVDKYTKDNKKNKMYNDIDEVKNKEANV-----YLKEKSKCKOVNFD 720

```
Db 485 DQWTRKKNQWDLNFKPISVKNAB-----KVQTAGIVTPYDILKQELDEFNVAFE 536
Qy 721 DKIFNESPNVEDYCKCKDEKYLNEIKYPTKHDIYDIDTFSTFG-DGTPI----- 772
Db 537 NEI-NKRDGAYIELC-----VCSVEEAK-KNTQEVVTVNDVNAKSOATNSNPIQPDVSS 589
Qy 773 -----SINANINEQSGKDTNIGNSSETSD-----SPVSHE-----PESDAAINV 812
Db 590 KAEKVPDGSHTGNV---SGQDSITTKAVTGDQNGNQTPAESDVQESDIAESVSAKV 646
Qy 813 E-----KLSGD-----BSSSETRGILINDPVSNTNNVNEVHDASNTGGSV----- 852
Db 647 DPQKSVSKRSDTASVTGIAEAGKLENLGAENSRPSESTVEANSPCDDTVNSASIPVVSGE 706
Qy 853 -----SNTSDITNGHSSSLNRTNAQDIKIGRSG-----NEQSDNQENSSHSS 896
Db 707 NPLVTPYNGLRHRSKONSDSGPAESMANPDSNK-----GETGKGQDNMDAKATKDSNSS 762
Qy 897 DNGSLT-----IGQVPSEDTQNTYDSQNPHRDTPNALASLPDSDDKINEIEG 944
Db 763 DGTSSATGDTTDAVDREINKVPEDRDKTVGSKDGGEDNSANKDAATVUGEDIRENSA 822
Qy 945 PDSSRD-----SENG-----RGDTTN---THDVRRTIV 971
Db 823 GGSTNDRSKNDTEKNGASTPDSKQSEDATLSKTESLESTESGDRTTNDTNSLENKNG 882
Qy 972 SERRVNSHDFTRNGMANNNAHOVIT-----QIENNGIIRGOBESAGNSVNYKDNPK 1023
Db 883 KBDLQXDFKSNTPNEBPSNDQTTDABGHRDSIKND---KAERKHMKNKOTFTKNTN 939
Qy 1024 RSNFSEN---DHKNQIOEYNSRDTKVRREIILKLSKONKNNEYSMEYCTYSBERNSSP 1080
Db 940 SHLNSNNLSNGKLDIKEYYRDVATREDIILMSSVRKCNNISLEYCNSVEDKISS- 998
Qy 1081 GPCSRERERKCCOISYCLFYFYFYSLEYNYCNISKEIKSPYKCPKSEG-QSSIPYFAA 1139
Db 999 NTCRSREKSNLCCSIDPCLNYFDVYSYLSKMGKGFEDPSYKFTKGGPKDKTYFAA 1058
Qy 1140 GGILVIVILLSSASRMKSNHEEDIGESNEATPEE-NYVNLKLSRI-----FNQEOVE 1193
Db 1059 AGAL--LILLIASRKMKNDS-----EATFNEFEYCDNIHRIPLMNNIEHMOP 1108
Qy 1194 TNISDYS 1200
Db 1109 STPLDYS 1115
```

RESULT 12

```
US-09-210-288-2
; Sequence 2, Application US/09210288
; Patent No. 6392026
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Welles, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/210,288
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fuller, Michael
; REGISTRATION NUMBER: 36,516
; REFERENCE/DOCKET NUMBER: NIH121.1FWDV1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1115 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium vivax
; US-09-210-288-2
```

```
Query Match 11.6%; Score 754; DB 2; Length 1115;
Best Local Similarity 24.5%; Pred. No. 2.5e-43;
Matches 296; Conservative 177; Mismatches 416; Indels 318; Gaps 57;
```

```
Qy 199 FKEIFLISVTEAKLLVKNNEGKDPISFCNELRNSPFSRSPFGDDMPGGTDRVKGY 258
Db 22 YKDFPSIT-----LINVHEGKYLILKR-----KLEKANNRDCVNF 59
Qy 259 IN-TKFSDDYKKEKVEKLNINKKEWE-KKANLWNIHNVHKNISKECAIIPAEBPQI 316
Db 60 LHSQVNNVLLERTIETILLECKNEYVKGNYL-----AKGH-----HC-----VEEDNL 105
Qy 317 NLMIKWNENFLMEKKRLFLNIDKCVENKRYEACFGGCLPCCS-SYTSFMKSKTOMEV 375
Db 106 ERWLQGTNER-----RSENIKYK--YGVTELKIKYAQWNGKRSSRLKES 149
Qy 376 L-----TNLYKQNSGVDKNNFLNDLFFKQNNQNDLDDFFKNEKEYDDLCDCRY----- 423
Db 150 IYGAHNFPGNSYMEKGOGDKTG-----BEKQGEHKTDKTDNGKGANNLVMLDYETSSN 204
Qy 424 -----TATIISKLNGPAKN-----DVD-----IASQINVNDLRGF-----G 455
Db 205 GQPAGTLDNLVLEFVTGHEGNSRKNSSNGNPYDIDHKKTISSAI-IN-HAFLQNTVMKN 261
Qy 456 CNYK-SNNEKSWNCTGFTTNKFPCTCEPPRRQTLCLGRTYLLHRGHEEDY-----KEH 507
Db 262 CNYKRRERDWDG-----NTKQDVCIIPDRYQOLCMKELTNLVNNTDTNFRDITFRKLY 316
Qy 508 LLGASIYEA-----OLLKYKYEKENALCSIIQNSYADLADIIGSDIHKDYGKK 559
Db 317 LKREKLIYDAAVEGDLILKANNRY-NKD-----FCKDIRWSLGDGFDIIMGTDMEGIGYSKV 372
Qy 560 MEENLKV-NKDKKNEESLKIREFKWDENKENVKMSAVLNK-----ETCKDYDK 612
Db 373 VENNLRSIFGTDEKAQOR-----RKQWNESKQIWTAMYSVVRKRLKGNPIWTKLNVA 427
Qy 613 FQKIPQFLRPFKEWGDGDFCEKREKIYSPESFKVEK-----KKQDCDENTCKNKGSEY 665
Db 428 VITEPQIYRWIREGRDYVSELPTVEV---QKLKEKCDGKINYTDKVKYKPPCQACKSY 484
Qy 666 KKWIDLKSEYKQVDKYTKDKNNKQMDYNDDEVNKEANV-----YLKESKECKQVNF 720
Db 485 DQWITRKKQWDLNFKPISVKNAB-----KVQTAGIVTPYDILKQELDEFNVAFE 536
Qy 721 DKIFNESPNVEDYCKCKDEKYLNEIKYPTKHDIYDIDTFSTFG-DGTPI----- 772
Db 537 NEI-NKRDGAYIELC-----VCSVEEAK-KNTQEVVTVNDVNAKSOATNSNPIQPDVSS 589
Qy 773 -----SINANINEQSGKDTNIGNSSETSD-----SPVSHE-----PESDAAINV 812
```

Db 590 KAEPVPGDSTHGVN---SQDSTTGKAVTGCQNGNQTPAESDVORSIAESVSAKV 646
Qy 813 E-----KLSGD-----ESSSETRGILDINDPSTVNNVNEVHDASNTQGSV-----852
Db 647 DPQKSVSKRSDDTASVTGIAEAGKENGAGNSRPSBSTVEANGSPGDDTVNSASIPVVSGE 706
Qy 853 -----SNTSDITGHGESSLRNTNAQDIKIRSG-----NEQSDNQESSHSS 896
Db 707 NPLVTPYNGLRHSKNSDSDGPAESMANPDSNK-----GETGKGQDNDAKATKDSNSS 762
Qy 897 DNGSGLT-----ICQVPSEDNTQNTYDSQPHRDTNPALASLPDDDKINEIEG 944
Db 763 DGTSSATGDTTDAVDREINKGVPEDRDKTVGSKDGGEDSANKDAATVVGEDRIRENSA 822
Qy 945 FDSRD-----SENG-----RGDTTSN---THDVRRTNIV 971
Db 823 GGSTNDRSKNDTEKNGASTPDSQSDATALSKTESLESTESGDRITNDTNSLEKNKG 882
Qy 972 SERVNSHDFIRGMANNAHQVIT-----QIENNGIIRGOEBSAGNSVNYKDNPK 1023
Db 883 KEKDLQKHDPKSNDDTNEPNSPQTTDAEGHDRDSIKND---KAERRKHMKDKDTFTKNTN 939
Qy 1024 RSNFSSN---DHKNIOEYNSRDKRVREIILKSKQNKCNNEYSMEYCTYSDERNSSP 1080
Db 940 SHLNSNNLSNGKLDIKEKYRDVKATREDIILMSSVRKNNNISLEYCNVEDKLSS - 998
Qy 1081 GPCSRERKKLCCQISDYCLKYFNFYSIYNCIKSEIKSPYCKFKSEG-QSSSIPYFAA 1139
Db 999 NTSREKSKNLCCSISDFCLNYPDVSVEYLSCKMKEFEDPSYKCTPKGFKDKTYFAA 1058
Qy 1140 GGLVVLVILLSSASRKGNEEYDICESNIEATFEE-NNYLKLKLSRI-----FNOEVOE 1193
Db 1059 AGAL--LILLIASRKMKNDSE-----EATNFEFYCDNIHRIPLMPNNIHQMP 1108
Qy 1194 TNISDYS 1200
Db 1109 STPLDYS 1115

RESULT 13

US-10-153-273-2

; Sequence 2, Application US/10153273

; Patent No. 6962987

; GENERAL INFORMATION:

APPLICANT: Sim, Kim L.
Chitnis, Chetan
Miller, Louis H.
Peterson, David S.
Su, Xin-zhaun
Welliams, Thomas E.

; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

; AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

; NUMBER OF SEQUENCES: 37

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Knobbe Martens Olson & Bear

; STREET: 620 Newport Center Drive 16th Floor

; CITY: Newport Beach

; STATE: California

; COUNTRY: US

; ZIP: 92660

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/153,273

; FILING DATE: 21-May-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/09/210,288

; FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael
REGISTRATION NUMBER: 36,516
REFERENCE/DOCKET NUMBER: NIH121.1FWDV1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1115 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium vivax
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-153-273-2

Query Match 11.6%; Score 754; DB 2; Length 1115;
Best Local Similarity 24.5%; Pred. No. 2.5e-43;
Matches 296; Conservative 177; Mismatches 416; Indels 318; Gaps 57;

Qy 199 FKEIFLISVNTAEKLLYNKEGKDPISFCNELNSPDSFRSSFIGDDMPFGGNTDVKGY 258
Db 22 YKDDFSIT-----LNYHEGKKYLIILKR-----KLEKANNRDVCNFF 59
Qy 259 IN--TKPSDYVKEQVKNLANNIKKEWWE-KNKANLWNHMIWNHKNISKECALIPAEEPQI 316
Db 60 LHPSQVNNVLLERTITLLECKNEYVKGNGYKL-----AKGH---HC-----VEEDNL 105
Qy 317 NLWIKENWENFLMEKRLFLNFKCVENKYEACFGGCRLPSC-SYTSFPMKSKTKQMEV 375
Db 106 ERWLQGTNER-----RSEENIKYK--YGVTELKIKYAQMNGKRSRILKES 149
Qy 376 L-----TNLYKKNSGVDKQNFNLDFKQNNKNDLDDFFKNEKEYDDLCDRY-----423
Db 150 IYGAHNFNGSNYMEGKGGDKTG-----BEKGEHKTDSKTDNGKANNLVMLDYETSSN 204
Qy 424 -----TATIIGKFLNGPAKN-----DVD-----IASQINVDNLRGF-----G 455
Db 205 GQPAGTLDNLVLEFVTGHEGNSRKNSSNGGNPYDIDHKTTISSAI-IN--HAFQNTVMKN 261
Qy 456 CNYK-SNNEKSNWCTGTFTNKPFGTCEPPRRQTLCLGRTYLLHRGHEEDY-----KEH 507
Db 262 CNYKRKRERDWDG-----NTKKDVCIPDRRYQLCMKELTNLVNNTDTNPHRDITFRKLY 316
Qy 508 LLGASIVEA-----QLLKYKKEKENALCSIIQNSVADLADIIGSDIIDKYGKK 559
Db 317 LKRKLIYDAAVEGDLKLNNRY-NKD---FCKDIRWSLGDGFDIIMGTDMEGIGYSKV 372
Qy 560 MEENLKNV-NKDKRNEESLKIIFREKWDENKENVKWSAVLKNK-----ETCKDYDK 612
Db 373 VENNLRSIFGTDEKAQR-----RKQWNEKSAQIWTAMMYSVKRKLKGNFICKLVA 427
Qy 613 FQKIPQFLRWFKWGDGDFCEKREKIKYSPESFKVEK-----KKKDCDBNTCKNKSEY 665
Db 428 VNIEPQIYRWIREWGRDYVSELPTVEV---QKLEKCDGKINYTDKKVKCVPPQONACKSY 484
Qy 666 KKWIDLKXSEYEQVDKYTKDNKNKQYDNIDEVKNEANV-----VLKESKECKOVNFD 720
Db 485 DQWITRKKQWDLVLSNKFISVKNAB-----KVQTAGIVTVPYDILKQELDSFNEVAPE 536
Qy 721 DKIFNESPEYEDMCKKCBDEIKYLNBEIKYPKTKHDYIDIDTFSDTFG-DGTPI-----772
Db 537 NEI-NKRDGAYIELC-----VCSVEEAK-KNTQEVTVNVDNAKSAQTSNPNISQPVDS 589
Qy 773 -----SINANINEQQSGKDTNTGNSETSD-----SPVSHE-----PESDAAINV 812
Db 590 KAEKVPDSTHGNVN---SQDSSITTGKAVTGCQNGNQTPAESDVORSIAESVSAKV 646
Qy 813 E-----KLSGD-----ESSSETRGILDINDPSTVNNVNEVHDASNTQGSV-----852


```
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,459A
; FILING DATE: 07-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israeleen, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2710 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
; US-08-568-459A-12

Query Match 7.68; Score 495.5; DB 1; Length 2710;
Best Local Similarity 19.5%; Pred. No. 6.5e-25;
Matches 268; Conservative 170; Mismatches 482; Indels 453; Gaps 57;

QY 112 LSENEGSGTNNNFANTSEISIGKDNKVTFIQKTRHLPACGIG-RKSIKWLRENSE 170
DB 912 INENYSDSRNCGPCT-----GKGDH-----GGVRWIGTEWSNIEGKK 953
QY 171 KIT---VCVPDRKIQLCVANFLNSRLFTM---EKFKEIFLISVNTAKL----- 213
DB 954 QTSKYNFLPDRRHMTCSNLENLDVGSVTKNDKASHLLGDVQLAAKTDAAEIIKRYKD 1013
QY 214 -----LYNKEGDKPSIFCNELNRSFDFRSPFSGDDM--DFGNTDVRVGYI----- 259
DB 1014 QNNIQLTDP IQQDQEAACRAVRYSFADLGDIIIRGRMDWDEKSSDTMETRLITVFNKIK 1073
QY 260 -----NTKFSDDYKKEKVEKLNKIKKEWKEKQKANKLNWNIWVNHKGNISKECALIPA 311
DB 1074 EKHDGIKNDPKYTGDESKKPAYK--KLRAWWEANRHOVWRAMKCATKGI-----CPGMPV 1128
QY 312 EE--PQNLMTKEWENFLMEKKRLFLNIIKDKCVE--NKKYEAC-----FGCRLPCSS 361
DB 1129 DDVIPQRLRWTEWAENWYCAQSOEYDKLKKIACDCMSKGDGKCTQGDVDCGKCAACDK 1188
QY 362 YTSFMKSKTQMEVLNLY-----KKKNSGVDKN-----NFLNDLPK--- 398
DB 1189 YKEIEKWEQWRKISKDYNLLYLQAKTTSTNPGRTVLGDDDPDYQOMVDPLTPIHKASI 1248
QY 399 ----- 398
DB 1249 AARVLVKRAAGSPTEIAAAPTSTYTAAGYIHOEIGYGCQEQTFQCEKKGATSTSTT 1308
QY 399 KNNKNDLDDFPNKEKYDDLCDCRYTATIKSFLNGPAKNDVDIASINVND--LRFGF- 455
DB 1309 KENKE--YTFKQPPEYATACD-----INRSQTEPKKEENVESACKIVEKILEGKNG 1361
QY 456 -----CNYKSNKEKSNCTGTFNKPFGTCPEPRRTQLCLGRYLLHRHBEEDYK--EH 507
DB 1362 RTTVGECNPK-ESYPDWDCKNNIDISHDGAECMPPRQKLC--YYIAHESQTENIKTDDN 1418
QY 508 LLGASIVAEA---QLLKYYKKEKBNALCSIIQN-----SYADLADLIKSGSDI 551
DB 1419 LKQAFIKTAAEFTLSQYKSKNDSEAKILDRGLIPSLRSMWYTFGDRICLNTDI 1478
QY 552 IKDYIG-KQWENLNK-VNKDKKRNESLSKIFREKWDENKENVWVMSAVL-----X 602
DB 1479 SKQNDVAKADKIGKFKFSKDGSKSPSGLS--ROEWKNTNGPEIWKGMCLCALTKYVTDTD 1536
QY 603 NKETCKD---YDK-----FQKIQPLRWFKNGDDDFCEKRKEKIVSFESFKV 646
```

```
DB 1537 NKRKIKNDYSYDKVNQSONGNFSLSEFAAKPQPLRWMIEMGEFCAERQKK---ENIIKD 1593
QY 647 ECKK----KDCDE--NTCKNKSEYKKWIDLKSEYKQVDKYT-----KDKNKQMYD 693
DB 1594 ACNEINSTQCCNDAKHRCNQACEAYQYEVYENKKKFSGGTNNFVLKANVQPDPEYKGYE 1653
QY 694 NIDEVKNKEANVYLKEKSECKDVNFDKIFNESPNE-----YEDMKCKDEIKYL 744
DB 1654 YKGVQPIQGNAYLLQCDNNKCSMDGNVLSVSPKEKPGKVAHYKPEKC-DCYQKGHV 1712
QY 745 NEIKYPK-----TKHDIYDIDTFSDT-----FGDGTPIISINANINEQ 781
DB 1713 PSIPPPPPVQPOPEAPTVTVVCSIVKTLFKDTNNFSDACGLKYGKTAPSSWKCPSPSD 1772
QY 782 QSGKOTSNKGSTSDSPV-----SHEPESD----- 807
DB 1773 KSGAG-ATTGKSGSDSGSICIPRRRLRYLVGKLQEWATALPQEGGAAPSHSRADDLNNAF 1831
QY 808 ---AAINV-----EKLSGDESSSETRGIL-----DINDPSVTNNVNEVHDASNTQ 849
DB 1832 IQSAAIETFTFLMDRYKEKKPGQDGSQALSQLTSTVSDDEEDPP-----DKLLQN 1882
QY 850 GSV-----SNTSDI--TNGHSESSLNRTTNAQDIKIGRSNEQ 885
DB 1883 GKIPDPFLRLMFYTLGDRDILVHGNTSDSGNTNG-----SNNNNIVLEASGNKE 1933
QY 886 SDN--QENSSHSDNSGSLTI-----GOVPSE--DNTON-----TYDSQNPHERDT 926
DB 1934 DMQIOEKEIQLPKNGGTPLPKSSAQTPDKWNEHAESIWKGMICALTYTEKNPD TSA 1993
QY 927 PNALASLPDDDKINEIEGFDSSRDSRSDGDTTSTNTHDVRRTNIVSERVNSHDFIRNGM 986
DB 1994 RGDENKIEKXDEV--YEKFFGSGTADKHGTASTPTGT----- 2027
QY 987 ANNAHQAIVITQIENNGIIRGOEESAGNSVYKDNPKRNFSSSENHKKNIQEYNSRDTK 1046
DB 2028 -----YKQYDYKVKLEDTSGAKTPSASDDTLPLSLDFVLRPPYFRYLEWGMQNFCK 2079
QY 1047 RVREEIIKLSKONK-----CNNEYSMEYCTYSD-ERNSSPGP 1082
DB 2080 KRKGKLAQIKHECKVEENGSGGSRGGITRQYSGDGEACHEMLPKNDGTVPDLEKPSCAKP 2139
QY 1083 CSR-----BERKKLCCQISDYCLKYENFYFYSIEYNNCIKSEIKSPY 1123
DB 2140 CSSYRKWIBSKGKEFEKQKAVEQQKDKCVNGSNKHDNGFCETLTSSKAKDF 2192
```

RESULT 18

```
US-08-487-826B-12
; Sequence 12, Application US/08487826B
; Patent No. 5993827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,826B

FILING DATE: 10-SEP-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Israelsen, Ned

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER: NIH121.001CP1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 2710 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORIGINAL SOURCE:

ORGANISM: Plasmodium falciparum

US-08-487-826B-12

Query Match 7.6%; Score 495.5; DB 1; Length 2710;

Best Local Similarity 19.5%; Pred. No. 6.5e-25;

Matches 268; Conservative 170; Mismatches 482; Indels 453; Gaps 57;

QY 112 LSEHDSNGTNSNNFANTSISIGKDNKQYTFQKTHLPACGKIK-RKSIKWTCRENSE 170
DB 912 INENYSDNRSGSGPCT-----GKGDH-----GGVRMRIGTEWSNIEGKK 953
QY 171 KIT---VCVPRKIQLCVANFLNRLTETM---EKFKEIFLISVNTAKL----- 213
DB 954 QTSYKNVFLPRRHEMCTSNLENLDVGSVTKNDKASHLLGDVLAQKTDAAEIIKRYKD 1013
QY 214 -----LYNKNGKDPSPICNBLRNSFDSFSSPFGDDM--DFGNGTDRVKGYI----- 259
DB 1014 QNNIQLTDPIQKQKQAMCRAVRYSFADLGDIIIRGRDMDWEDKSSDTMETRLITVFKNIK 1073
QY 260 -----NTKFSYDYKEKVEKNLNNIKKEWEKQKANKLNNHMIVNHKGNISKECALIPA 311
DB 1074 EKHDIKDNPKYTGDESKKPAYK--KLADWWEANRHQVWRAMKCATKGII---CPGMPV 1128
QY 312 EE--POINLWIKWENFLMEKKELFLNLIKDKVE--NKKYEAC-----PGCRLPCSS 361
DB 1129 DDYIPQRLRWMTWAEWYCKAQSQR-----DKLKKICADKWSKDGKCTQGDVDGCKKAAADK 1188
QY 362 YTSFMKSKTKQMEVLTNLY-----KKNNSGVVDKN-----NPLNDLFK--- 398
DB 1189 YKEIEKWNEQWRKISDKYNLLYLQAKTTSNPGRTVLGDDDDPDYQQWVDPLTPIHKASI 1248
QY 399 ----- 398
DB 1249 AARVLVKRAAGSPTEIAAAAPITPYSTAAGYIHOEIGYGCQEQTFQCEKKGHGTATSTTT 1308
QY 399 KNNKNDLDDFPKEKEVDLDCRYTATI KSFLLNGPAKNDVDIASQINVD--LRGFG- 455
DB 1309 KENKE--YTFQKPPPEYATACD-----INRSQTEPKKKEENVESACKIVEKILEGKNG 1361
QY 456 -----CNYSNNKSNKWNCTFTNPKFPGTCPPRROTLCIGRYTLHLRHGHEDYK--EH 507
DB 1362 RTTVGECNPK--ESYPDWDCKNNIDI SHDGACMPRRQKQLCL--YYIAHESQTENIKTTDN 1418
QY 508 LLGASIVEA---QLLKVKYKKEBNALCSIIQN-----SYADLADIIGSDI 551
DB 1419 LKDAFIKTAATAETFLSWOYKSKNDSEAKILDRGLIPSOFLRSMWMTFGDYRDI CLNTDI 1478
QY 552 IKDYVG--KWEENLANK-VNKDKKNEBSLKIIFREKWNENKENVKWSAVL-----K 602
DB 1479 SKQNDVAKAKDKTGKFFSKDGSKSPGLS--ROEWKTKNGPEIWKGLMCLTKYVTDTD 1536
QY 603 NKETCKD---YDK-----FQKIPQFLRWFKWGDGDFCEKREKIKIYSFESFKV 646

DB 1537 NKRIKNDYSDYKVNQSGNPSLEEAFAKQFLRWLEWGEERFAERQKK---ENIKD 1593
QY 647 ECKK-----KDCDE--NTCKKCKSEYKKWIDLKSGEYKQVDKYT-----KDKNKQYD 693
DB 1594 ACNEINSTQOCNDAKHRCNQACRAYQBYEVENKKKFGSGQTNNFVLKANVQDPPEYKGYE 1653
QY 694 NIDEVQKNEANVYLKESKECKQVNFDDKIPNESPNE-----YEDMCKKCDDEIKYL 744
DB 1654 YKGVQPIQNEVYLLQKCDNNKSCMDGNVLSVSPKPKFGYAHKYPEKC--DCYQGRHV 1712
QY 745 NEIKYPK-----TKHDIYDIDTFSDT-----FGDGTPIISINANINEQ 781
DB 1713 PSIPPPPPVQPOPEAPTIVTVVCSIVKTLTKDTNNFSDACGLKYGTAPSWKCI PSDT 1772
QY 782 QSGKDTNTGNSSETSDSPV-----SHEPESD----- 807
DB 1773 KSGAG-ATTGKSGSDSGSICIPRRRLRYVKGLEWATALPQGEAAPSRSRDLRNF 1831
QY 808 ---AAINV-----EKLSGDESSESTRGIL-----DINDPSVTNNVNVHDSNTQ 849
DB 1832 IQSAAIETFFLWDRYKEEKFPQDGSQOALSQLTSTYSDEEDPP-----DKLLQN 1882
QY 850 GSV-----SNTSDI--TNHSESSLNRTTNAQDIKIGRSGNEQ 885
DB 1883 GKIPDFLRLMFYTLGDYRDILVHGGNTSDSGNTNG-----SNNNNIVLEASGNKE 1933
QY 886 SDN--QENSHSSDNSGLTI-----CQVPE-----DNTQN-----TVDSONPHRDT 926
DB 1934 DMQIKQIKBIQILPKNGGTFLVPKSSAQTDPKWNEHAESIWKGMICALTYTEKNPDTSA 1993
QY 927 PNALASIPSDDKINEIEGFDSRDSSENGRGDTTNTHDVRRTNIVSERRVNSHDFIRNGM 986
DB 1994 RGDENKIEKDEV--YEKFGSTADKHGTASTPTGT----- 2027
QY 987 ANNAHGOYTQIENNGIIRGQESAGNSVNYKDKRSPSSNDHKKNTQEVNSRDTK 1046
DB 2028 -----YKQVDYEVKVLDTSGAKTPSASSDTPLLSDFVLFPFVFRYLEEWGQNFCK 2079
QY 1047 RVREIILKSKQNK-----CNNEYSMEYCTYSD--ERNSSPCP 1082
DB 2080 KRKHKLAAQIHECKVEENGSGRRGGITROYSGDGEACNEMLPKNDGTVLPLEKPSCAKP 2139
QY 1083 CSR-----BERKKLCCQISDYCLKYFNFYIEYVYICIKSEIKSPY 1123
DB 2140 CSSYRWKIESKGKFEKQKAYEQKQKCVNGSKNDNGFCETILTSSKANDP 2192

RESULT 19

US-09-210-288-12

Sequence 12, Application US/09210288

Patent No. 6392026

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.

APPLICANT: Chitnis, Chetan

APPLICANT: Miller, Louis H.

APPLICANT: Peterson, David S.

APPLICANT: Su, Xin-zhaun

APPLICANT: Wellem, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe Martens Olson & Bear

STREET: 620 Newport Center Drive 16th Floor

CITY: Newport Beach

STATE: California

COUNTRY: US

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS


```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/929,329
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Breen, John P
; REGISTRATION NUMBER: 38,833
; REFERENCE/DOCKET NUMBER: 835910-28685
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 231-7745
; TELEFAX: (317) 231-7433
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1507 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium yoelii
; US-08-929-329-5

Query Match 6.7%; Score 432; DB 2; Length 1507;
Best Local Similarity 19.6%; Pred. No. 6.9e-21;
Matches 261; Conservative 161; Mismatches 403; Indels 504; Gaps 55;

QY 33 NRODESSDISRV-----NSPELNHNKNTIYDSYEDVNNKL-----INSFVENKSVKK 81
DB 436 NQIIEYDIEENVKIWKRNSELSNLYNDKFKXNSINDKGMAMNDNDINEIKENSQLOT 495
QY 82 KRSLSFNNKTK-----SYDIIPPSVSRNDK-----FNSLSNEDNS-GNTNS--NNF 127
DB 496 NK-----GNETKTKYGLYNYPTIPSYLQIHQKMKELKNYMDSENSFTSFHNTNATHY 550
QY 128 ANTSEISGKDNKYOT-----ICRESEKTVCPDRKIQLCVANFNSRLB-TMEKFK 156
DB 551 EGNKSTGVNNKRENTYGTQDINLRNNNNYQPKNPQAEYMDRFDEKNIHY---- 606
QY 157 KRKSIAK-----ICRESEKTVCPDRKIQLCVANFNSRLB-TMEKFK 200
DB 607 ----IDKQDGKYGSKLYNIISHTADTIQSLLITDKODICPNHYSPOGACQCPNYG 662
QY 201 EIFLI-----SVNTEAKLLYNKQKPSIFCNELRNSFSDFRSSFIGDDMDFGNTDRVK 256
DB 663 KSIIVKTPESINGNEHL--NSN-----FLNEIRT----- 689
QY 257 GYINTKFSDYKEKQVE-----KLNNIKKEWKEKN---KANLWNH----- 293
DB 690 GYLNK-----YKSNVELPYEKSLGAMHGGDLSCPKSWDEENLYKKNRDYNDYMDCKSTV 744
QY 294 -----MIVNHK-----GNISKECAIIPAEBPQINLWKENNFPM 329
DB 745 MKSTIPLKMFEDYTKLLYFLGLYGLGRLGSLNSIKVNIKFSQPNNTIL-----PM 795
QY 330 EKKRLFLNIKD-----KCVNKKYEACFGGCRLPSCSYTSFMKSKTQMEVL---- 376
DB 796 FNPSSIKNLLDCSLYSYCLGPLENAYNNKCF-----RSLPAYFHETNECIIIGTHE 848
QY 377 ---TNLYKKKNSGVDKNNFLNDLPKKN-----NKNDLDDFFKNEKEYDDLCDCRYATI-- 427
DB 849 QERNNNCKTRSDTDKPCNQ--VRKNISTKNTVTVTSFIR--PDYBEKCPPRYPPLKPKS 904
QY 428 -----IKSFNGPAKNDVDI----- 442
DB 905 FGKYDEBTGCKGSLIN--KKNIINIPFSSCLEYFMIPSVLQRTKQNGWGVWVASES 962
QY 443 ASQINVDLRGFCNGYKNSKSNKNTGTFTNKFPGTCEPPRRQTLCLGRTYLLHRGHEE 502
DB 963 VNSSNLYNAGK-ECY--INEKP-NCVIDKYNHF-----SFTSLTTNDI 1002
```

RESULT 22

US-10-087-013-8

; Sequence 8, Application US/10087013

; Patent No. 6855323

; GENERAL INFORMATION:

; APPLICANT: Arthur Scherf

; APPLICANT: Louis H. Miller

; APPLICANT: Benoit Gamain

; APPLICANT: Dror I. Baruch

; APPLICANT: Pierre Buffet

; APPLICANT: Christine Scheidig

; APPLICANT: Jurg Gysin

; APPLICANT: Bruno Pouvelle

; APPLICANT: No. 685323utaka Fujii

; APPLICANT: Joseph Smith

; TITLE OF INVENTION: IDENTIFICATION OF THE DOMAIN OF

; FILE OF INVENTION: PLASMODIUM FALCIPARUM ERYTHROCYTE MEMBRANE PROTEIN 1

; FILE REFERENCE: NIH176.001C1

; CURRENT FILING DATE: 2002-02-21

; PRIOR APPLICATION NUMBER: US/10/087,013

; PRIOR APPLICATION NUMBER: PCT/US00/24195

; PRIOR FILING DATE: 2000-09-01


```
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 921 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
; US-09-210-288-8

Query Match          5.6%; Score 360.5; DB 2; Length 921;
Best Local Similarity 22.4%; Pred. No. 3e-16;
Matches 166; Conservative 83; Mismatches 231; Indels 261; Gaps 33;

Qy 217 KN-EGKPSIFCNELRNPSFDRSSFGDDMDPGNTDRVKGYNINTEKPSDYKKN-----271
Db 58 KNCNGGDTGKCG-SNGKKGNDYITASDIENGNS---IGNIDMVVSD--KDANGFNG 111
Qy 272 -----VEKLNKIKKEWKK-ANLWNHVIWNHKNISKECALIPAEPO---INLWIK 321
Db 112 LDACGSANIPKIRKEQKCAKVGCLDVCGLKNNGSIDK-----DQKQIIIRALLK 164
Qy 322 EWNENFLMEKKRLFLNIDKCKVENKYYACFGGCRPCSSYTSFMKSKTQMEVLTNLYK 381
Db 165 RWVEYFLDYNNKINAKI-SHCTKDNSTCTNDCPNKCTCVEEWINQKRTWKNIKKHYK 223
Qy 382 KNSGVDKN--NFLNDLF-----KKNK-----NDLDDFFPKKEKYD 416
Db 224 TONENGDNMKSLVTDILGALQPSQDVNKAIKPCSGLTAFESFCGLNGADNKEKGEYD 283
Qy 417 DL-----CDRYTATIISFLNG-----434
Db 284 DLVLCMLKNLEKIQEQCKKKHGETSVE---NGGKSCPTLDNTTLEEEPIEENQVEAPNI 340
Qy 435 -----PA-----KNDVDIAS 444
Db 341 CPKQTVEDKKEBEETCTPASPEKVPVHVRWRTTPPEVFKIWRGRNKTTCIEVA 400
Qy 445 QINVDNLG---FGNYSKNNEKSWNCTGTTNKP-----GTCBPPRRQTLC-----490
Db 401 EM-LKDNKNGRTTVGECYRKETYSEWTC---ESKIKMGQHGACIPPRRQKLCLEKIM 456
Qy 491 -----GRYLL-----HRGHEEDYKEHLGASIEAQLLYKYEKDE 528
Db 457 TTNELKVFAPKCAAAETFLWQNYKDKNGNAEDLDEKLGIIIPDFKQMFY-----511
Qy 529 NALCSIIQNSYADLADIIKGSIIKDYKGMENLANKV--NKDKRNEESLKI-FREKW 585
Db 512 -----TPADYEDICLGTDISK---KDTSGVGKVCNIDDVFKIISNIRYKSW 559
Qy 586 WDKENKVNKWSAVLK-----NKETCK-----DYKFKQIP 617
Db 560 WETNGPVWEGMLCALSVDTSLNINNVNPTHKLTGEGNNFEKVFSGDSSTLTKFSERP 619
Qy 618 QPLRWFKEGDDFCRKEKELYSPEKVECK-KDCDEN-----TKKKCSYK 666
Db 620 QFLRLTWGNGNFCKEK-----YKVLAKCKDCVDGDKNGKVCACKQCKQYH 673
Qy 667 KWLIDKKSEYKQVDKYTKDKKKMYDNIDVQNK-EANVYLKESKECK-----DWNF 719
Db 674 SWIGWIDNYKKQGRYTEVKKIPLYKEDKDKVNSDDARDLYLTQLQNMKCVNGTTDENC 733
Qy 720 DDKIFNESPNEVDKCKKDE 740
Db 734 EYKCMHKTSTNSDMPESLDE 754

RESULT 26
US-10-153-273-8
; Sequence 8, Application US/10153273
; Patent No. 6962987
```

```
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; Chitnis, Chetan
; Miller, Louis H.
; Peterson, David S.
; Su, Xin-zhaun
; Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/153.273
; FILING DATE: 21-May-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/210.288
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Fuller, Michael
; REGISTRATION NUMBER: 36.516
; REFERENCE/DOCKET NUMBER: NIH121.1FWDVI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 921 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-153-273-8
```

```
Query Match          5.6%; Score 360.5; DB 2; Length 921;
Best Local Similarity 22.4%; Pred. No. 3e-16;
Matches 166; Conservative 83; Mismatches 231; Indels 261; Gaps 33;

Qy 217 KN-EGKPSIFCNELRNPSFDRSSFGDDMDPGNTDRVKGYNINTEKPSDYKKN-----271
Db 58 KNCNGGDTGKCG-SNGKKGNDYITASDIENGNS---IGNIDMVVSD--KDANGFNG 111
Qy 272 -----VEKLNKIKKEWKK-ANLWNHVIWNHKNISKECALIPAEPO---INLWIK 321
Db 112 LDACGSANIPKIRKEQKCAKVGCLDVCGLKNNGSIDK-----DQKQIIIRALLK 164
Qy 322 EWNENFLMEKKRLFLNIDKCKVENKYYACFGGCRPCSSYTSFMKSKTQMEVLTNLYK 381
Db 165 RWVEYFLDYNNKINAKI-SHCTKDNSTCTNDCPNKCTCVEEWINQKRTWKNIKKHYK 223
Qy 382 KNSGVDKN--NFLNDLF-----KKNK-----NDLDDFFPKKEKYD 416
Db 224 TONENGDNMKSLVTDILGALQPSQDVNKAIKPCSGLTAFESFCGLNGADNKEKGEYD 283
Qy 417 DL-----CDRYTATIISFLNG-----434
Db 284 DLVLCMLKNLEKIQEQCKKKHGETSVE---NGGKSCPTLDNTTLEEEPIEENQVEAPNI 340
```



```
QY 1094 QISDYC-LKYFNFSYIEYNCIKSEIKSPYKCFKSEQSSIPYPAAGGILV 1144
Db 866 --KEACGLKYGP-----GGKEKPNKCVTPSGVSTATSGKGAICV 905

RESULT 28
US-10-087-013-7
; Sequence 7, Application US/10087013
; Patent No. 6855323
; GENERAL INFORMATION:
; APPLICANT: Arthur Scherf
; APPLICANT: Louis H. Miller
; APPLICANT: Benoit Gamain
; APPLICANT: Dror I. Baruch
; APPLICANT: Pierre Buffet
; APPLICANT: Christine Scheidig
; APPLICANT: Jurg Gysin
; APPLICANT: Bruno Pouvelle
; APPLICANT: No. 6855323utaka Fujii
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ERYTHROCYTE MEMBRANE PROTEIN 1
; TITLE OF INVENTION: (PFEMP1) THAT MEDIATES ADHESION TO CHONDROITIN SULFATE A
; FILE REFERENCE: NIH176.001C1
; CURRENT APPLICATION NUMBER: US/10/087,013
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: PCT/US00/24195
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/152,023
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-087-013-7

Query Match 5.4%; Score 348; DB 2; Length 294;
Best Local Similarity 32.9%; Pred. No. 4.3e-16;
Matches 94; Conservative 37; Mismatches 87; Indels 68; Gaps 15;

QY 456 CNYKSNNEKSWNC-TGFTNKPFGTCPPRRQTLCLGRTYLLHRGHEDYK-----EHL 508
Db 13 CNAKNRKNEQCDQNTFPVDCNEGVCMPRRKSICIHNLTL-----EQTKNKYQLREAF 67
QY 509 LGASIEYEAQLLYKYKEDEVALCSIIQN-----SYADLADIIGSDIIKDY 555
Db 68 IKCAAKETNLLWDKYK-NDKNEABELLKKKIPEDFMRIMFYTFGDFRDFCLENDMGKDV 126
QY 556 YGKMEENLKV-NKDKKRNEESLKI PREKWDENKENVKMSAVL-----KNKE 605
Db 127 --DKVKNINKVFNNSKRGFK--KIDPENWNNNGPOIWNGLCALIHADTKDSIRKND 182
QY 606 TKDQYDK-----FQKIPQFLRMFKWGGDDFCCKRKEKIYSPFSFKVEC 648
Db 183 NYK-YEKTILAKRDGSGNMTLSEFAKPKPLRWFWYDDYCKRQKYLTEVAS---TC 238
QY 649 KKGD-----CDNTCKNKCSEYKWKIDLKSEYKQVDKYTKDKNK 689
Db 239 KSIDGGQLKCDRG-CNNKCDSEYKWKIRKKBEWNLQ-DKYYKDKRE 282

RESULT 29
US-10-087-013-9
; Sequence 9, Application US/10087013
; Patent No. 6855323
; GENERAL INFORMATION:
; APPLICANT: Arthur Scherf
; APPLICANT: Louis H. Miller
; APPLICANT: Benoit Gamain
; APPLICANT: Dror I. Baruch
; APPLICANT: Pierre Buffet
```

```
; APPLICANT: Christine Scheidig
; APPLICANT: Jurg Gysin
; APPLICANT: Bruno Pouvelle
; APPLICANT: No. 6855323utaka Fujii
; APPLICANT: Joseph Smith
; TITLE OF INVENTION: IDENTIFICATION OF THE DOMAIN OF
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ERYTHROCYTE MEMBRANE PROTEIN 1
; FILE REFERENCE: NIH176.001C1
; CURRENT APPLICATION NUMBER: US/10/087,013
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: PCT/US00/24195
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/152,023
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-087-013-9

Query Match 5.1%; Score 333.5; DB 2; Length 351;
Best Local Similarity 27.9%; Pred. No. 5.6e-15;
Matches 92; Conservative 46; Mismatches 131; Indels 61; Gaps 11;

QY 442 IASQINVNDLRGFCGNYKSNNEKSWNCCTGTNTKFPCTCEPPRRQTLCLG-RTYLLHRGH 500
Db 9 ISEQIEKNH-----NCKKTEDAKWCKENTKLGDEGVCMPRRQNLCVHYLTTLNDDSK 64
QY 501 EEDYKEHLGASIEYEAQLLYKYKEK-----DENAL-----CSIQNSYADLADIIGK 548
Db 65 EEDLEAFIKSAAAEFLLRQYNSKNVEDDKILHRDWIPPEFSPMYTGGDYDCLD 124
QY 549 SDI---IKDY----YGKMBENLKNVKNKDNESLKI PREKWDENKENVK-----595
Db 125 TDISEKIADHVTAKKKITAVFQKIGSKTTNGKKVLE--REGWWEYGLSIWKMGLCAL 182
QY 596 -----VMSAVLKNKETCKDYDKFQKIPQFLRMFKWGGDDFCCKRKEKIY 639
Db 183 SYNTETKMDGVRTYLMKYIYKNDIKYELSEFASRPPPLRWVTEWGEDFVKRKKSLV 242
QY 640 SPFSFKVEC-----KKQCDEN-----TCNKNKSEYKWKIDLKSEYKQVDKYTKDKN 688
Db 243 SLKKKCDSCILRNNGTNGTKTCDNENCGACKTQCEKYKKNMERWKYHSSQKKKFLYKN 302
QY 689 KQYDNDIDEVKNKEANVYLKESKECKDVN 718
Db 303 SATYNN--GLAVKEANSETYKNDPEVTEAN 330

RESULT 30
US-10-087-013-11
; Sequence 11, Application US/10087013
; Patent No. 6855323
; GENERAL INFORMATION:
; APPLICANT: Arthur Scherf
; APPLICANT: Louis H. Miller
; APPLICANT: Benoit Gamain
; APPLICANT: Dror I. Baruch
; APPLICANT: Pierre Buffet
; APPLICANT: Christine Scheidig
; APPLICANT: Jurg Gysin
; APPLICANT: Bruno Pouvelle
; APPLICANT: No. 6855323utaka Fujii
; APPLICANT: Joseph Smith
; TITLE OF INVENTION: IDENTIFICATION OF THE DOMAIN OF
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ERYTHROCYTE MEMBRANE PROTEIN 1
; FILE REFERENCE: NIH176.001C1
; CURRENT APPLICATION NUMBER: US/10/087,013
; CURRENT FILING DATE: 2002-02-21
```

```
; PRIOR APPLICATION NUMBER: PCT/US00/24195
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/152,023
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-087-013-11

Query Match          5.1%; Score 327.5; DB 2; Length 308;
Best Local Similarity 27.8%; Pred. No. 1.2e-14;
Matches 85; Conservative 46; Mismatches 118; Indels 57; Gaps 10;

Qy 455 GCNYKSNNEKSWNCTGTTNKPFCPTCEPPRRQTLCGLRTYLLHRGHE-----EDYKEHLG 510
Db ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 511 ASIYEAOQLLYKYEKDENALCISI-----TONSYADLADIIGSDIIOYK 558
Db : ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 66 CAAKETHFAWHYKKNVNAENELKSGKIPBGFQKQMYTTFGDPDRDIFFGTDISSCRYIK 125
Db : ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 559 KMEENLKNVKNKKRNE-----ESLKIPEKRWNDENKENVKMSAVLKN-----K 604
Db : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 126 DTSQTIKSLGDOATTEKGDTHIDNKKL--QEWWTTHGPKIWEGLCALTNGLSESEK 183
Qy 605 ETCKDY-----DKFQKIPQFLRWPKWGGDDFCCKRKEKIYSFSEKVECK 649
Db : ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 184 NILQDYSYNLNAEKDDCCLEKFAKQFLRWYVWSDFCRERKKLEKVEDVCIAK 243
Qy 650 -----KKDCDENTCNKCKSEYKWIIDLKSEYEQVDKYTKDNKNKMYNDIDEVKNKEAN 704
Db : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 244 DYEGCKNKNNSCVKCKEYENITGKTIQYESQEGKFNTERKQKK-PEYNSYSKKDAS 302
Qy 705 VYLEK 710
Db ||| | | |
Qy 303 EYLAKD 308

RESULT 31
US-08-568-459A-10
; Sequence 10, Application US/08568459A
; Patent No. 5849306
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,459A
; FILING DATE: 07-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned

; PRIOR APPLICATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CPI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 700 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
US-08-568-459A-10

Query Match          5.0%; Score 325; DB 1; Length 700;
Best Local Similarity 24.4%; Pred. No. 5.8e-14;
Matches 120; Conservative 70; Mismatches 161; Indels 140; Gaps 22;

Qy 478 GTCEPPRRQTLC-----LGRTYLLHRGHEEDYKEHLGLGASIVYEAQLLYKY-----KE 525
Db ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 526 KBNALCISIIONSADLADIIGSDIIOY-----YKQWEENL-----NKVAKDKRNE 575
Db : ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 63 DSASQLCTVLARSFADIGDIVRGKLYGVDNKEKEQRKLEKLDIFPKIHKDVNKTN 122
Qy 576 ESLKIP-----REKWDENKENVVK-----VMSAVLKNKET-----K 606
Db : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 123 GAQERYIDDAKGGDFPQLREDWMTSNRETVMKALI CHAPKEANYFKTACNVGKGTNGQC 182
Qy 607 -CKDYD---KFQKIPQFLRWPKWGGDDFCCKRKEKIYSFSEKVECK-----K 651
Db : ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 183 HCLGGDVPYTFDVPQYLAWFBEWAEDFCRKKKCL---ENLQCRDYEQNLYCSNGY 239
Qy 652 DCDE-----NTCKNCKSEYKWIIDLKSEYEQVDKYTKD-----KNK 689
Db : ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 240 DCTKTIYKKGKLVIGSHCTNCSVMCRMYETWIDNQKKEFLKQKRYETETISGGSGSKSPK 299
Qy 690 KMYDNIDEVKNKEANYLVLEKSEKEDVNFDD-----KIFNESPNEVEDMC-----K 736
Db : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 300 RTRKAARSSSSDDNGYESKFKYKGLKEVGQYQVDKFLKILNK-----EGICQKQPOVGNE 354
Qy 737 KCDEIKYLAIEIKYPKT--KHDIYDIDTFSDFDGTGPISINANINEQOQSGKDSNTGSE 794
Db : | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 355 KADNVDFTE-KVKTFSRTEICEPCPCWGLEKGGPPWKV-----KGDKTCGSAKTK 405
Qy 795 T-----SDSPVSHPEPESDAAINVEKLSGDESSESTRGIL-----DINDPSVTNNVNE 841
Db : | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 406 TYDPKNITDIPVLYPKSQNLKYNKFCCKGAPGGQIKKQWCYCYDEHRPSSKNNNC 465
Qy 842 VHDASN--TQG 850
Db : | : | | |
Qy 466 VEGTWDKFTQG 476

RESULT 32
US-08-487-826B-10
; Sequence 10, Application US/08487826B
; Patent No. 5993827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
```

ADDRESSEE: Knobbe Martens Olson & Bear
 STREET: 620 Newport Center Drive 16th Floor
 CITY: Newport Beach
 STATE: California
 COUNTRY: US
 ZIP: 92660
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/487.826B
 FILING DATE: 10-SEP-1993
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Israelsen, Ned
 REGISTRATION NUMBER: 29,655
 REFERENCE/DOCKET NUMBER: NIH121.001CP1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 235-8550
 TELEFAX: (619) 235-0176
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 700 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Plasmodium falciparum
 US-08-487-826B-10

Query Match 5.0%; Score 325; DB 1; Length 700;
 Best Local Similarity 24.4%; Pred. No. 5.8e-14;
 Matches 120; Conservative 70; Mismatches 161; Indels 140; Gaps 22;

Qy	478	GTCEPPRQTL	-----LGRTYLLHRCHEEDYKEHLGASIEAQLLKYY-----KE 525
Db	8	GACAPYRLHCDYNLESIDTSTTHKLLLE-----VMAAKYEGNSINTHYTOHRTNE 62	
Qy	526	KDENALCSIIQNSYADLADIIGSDIHKDY-----YKQMEENL-----NKNVKDKKRN 575	
Db	63	DSASQLCTVLARSADIGDIVRGKDLVGYDNKEKEQKLEQKLDIFKKIHKDVMTN 122	
Qy	576	ESLKIF-----REKWDENKENVK-----VMSAVLNKKT-----K 606	
Db	123	GAQERYIDDAKGGDFQLREDWMTSNRETVMKALICHAPKEANYFIKTACNVGKTNGQC 182	
Qy	607	-CKDYD---KFQKIPQFLRFKKGDDFCEKREKIYFESFKVECK-----K 651	
Db	183	HCIGGDVPTFYDYVYQYLRWFEEWAEDFCRKKKKL-----ENLQKCRDYEQNLCSNGY 239	
Qy	652	DCDE-----NTCKNCKSEYKKWIDLKSEYKQVDKYTKD-----KNK 689	
Db	240	DCTIYKKGKLVIGEHCNCSVWCRVETWIDNQKKEFLQKRYETISGGSGSKPK 299	
Qy	690	KMYDNIDEVKNKEANVYLKEKCKQVNFDD-----KIFNESPEYEDMC-----K 736	
Db	300	RTKAARSSSSDDNGYESKFYKLVGYQVDKFLKLNK-----EGICQKQPOVQNE 354	
Qy	737	KCDRIKYLNEIKYPT--KHDIYDIDTDFGDTGISINANIEQSGKDTNTGNSE 794	
Db	355	KADNVDPFNE--KYVKTFSRTICEPCPWCGLKGGPPMKV-----KGDKTCSAKTK 405	
Qy	795	T-----SDSPVSHEPESDAINVEKLSGSSSETRGIL-----DINDPVSNTNNE 841	
Db	406	TYDPKNITDIPVLPDKSQQNLKYYKVFCEKAGPGGQIKKWQCYIDEHRPSSNNNC 465	
Qy	842	VHDASN--TQG 850	
Db	466	VEGTWDKFTQG 476	

RESULT 33
 US-09-210-288-10
 Sequence 10, Application US/09210288
 Patent No. 6392026
 GENERAL INFORMATION:
 APPLICANT: Sim, Kim L.
 APPLICANT: Chitnis, Chetan
 APPLICANT: Miller, Louis H.
 APPLICANT: Peterson, David S.
 APPLICANT: Su, Xin-zhaun
 APPLICANT: Wellem, Thomas E.
 TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
 AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
 NUMBER OF SEQUENCES: 37
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Knobbe Martens Olson & Bear
 STREET: 620 Newport Center Drive 16th Floor
 CITY: Newport Beach
 STATE: California
 COUNTRY: US
 ZIP: 92660
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/210,288
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Fuller, Michael
 REGISTRATION NUMBER: 36,516
 REFERENCE/DOCKET NUMBER: NIH121.1FWDV1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 235-8550
 TELEFAX: (619) 235-0176
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 700 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Plasmodium falciparum
 US-09-210-288-10

Query Match 5.0%; Score 325; DB 2; Length 700;
 Best Local Similarity 24.4%; Pred. No. 5.8e-14;
 Matches 120; Conservative 70; Mismatches 161; Indels 140; Gaps 22;

Qy	478	GTCEPPRQTL	-----LGRTYLLHRCHEEDYKEHLGASIEAQLLKYY-----KE 525
Db	8	GACAPYRLHCDYNLESIDTSTTHKLLLE-----VMAAKYEGNSINTHYTOHRTNE 62	
Qy	526	KDENALCSIIQNSYADLADIIGSDIHKDY-----YKQMEENL-----NKNVKDKKRN 575	
Db	63	DSASQLCTVLARSADIGDIVRGKDLVGYDNKEKEQKLEQKLDIFKKIHKDVMTN 122	
Qy	576	ESLKIF-----REKWDENKENVK-----VMSAVLNKKT-----K 606	
Db	123	GAQERYIDDAKGGDFQLREDWMTSNRETVMKALICHAPKEANYFIKTACNVGKTNGQC 182	
Qy	607	-CKDYD---KFQKIPQFLRFKKGDDFCEKREKIYFESFKVECK-----K 651	
Db	183	HCIGGDVPTFYDYVYQYLRWFEEWAEDFCRKKKKL-----ENLQKCRDYEQNLCSNGY 239	
Qy	652	DCDE-----NTCKNCKSEYKKWIDLKSEYKQVDKYTKD-----KNK 689	

Db 240 DCTKIYKKGKLVIGHECTNCVWCRMYETWIDNQKEFLKQKRKYETEISGGSGKSPK 299

Qy 690 KMYDNIDEVKNKEANVYLKESKECDVNFDD-----KIFNESPNYEYEDMC-----K 736

Db 300 RTKRAARSSSSDDNGYESKFYKLEKGVQDVDFKFLKLNK-----EGICQKQPQVGN 354

Qy 737 KCDEIKYLNIEIKYKPT--KHDIYDIDTFSDTFGDGTPIISANINAEQOQSGKDTSTNGSE 794

Db 355 KADNVDTNE-KYVKTFSTCEPCWGLEKGGPPWKV-----KGDKTCSAKTK 405

Qy 795 T-----SDSPVSHPESDAAINVEKLSGDESSSETRGIL-----DINDPVSVTNNVNE 841

Db 406 TYDPKNITDIPVLPDKSQNNILKYNFCCKGAPGGQIKKWQCYDEHRPSSKNNNC 465

Qy 842 VHDASN--TQG 850

Db 466 VEGTWKFTQG 476

RESULT 34

US-10-153-273-10

; Sequence 10, Application US/10153273

; Patent No. 6962987

; GENERAL INFORMATION:

; APPLICANT: Sim, Kim L.

; Chitnis, Chetan

; Miller, Louis H.

; Peterson, David S.

; Su, Xin-zhaun

; Wellens, Thomas E.

; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

; AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

; NUMBER OF SEQUENCES: 37

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Knobbe Martens Olson & Bear

; STREET: 620 Newport Center Drive 16th Floor

; CITY: Newport Beach

; STATE: California

; COUNTRY: US

; ZIP: 92660

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/153,273

; FILING DATE: 21-May-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/09/210,288

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Fuller, Michael

; REGISTRATION NUMBER: 36,516

; REFERENCE/DOCKET NUMBER: NIH121.1FWDV1

; TELECOMMUNICATION INFORMATION:

; TYPE: amino acid

; TELEPHONE: (619) 235-8550

; TELEFAX: (619) 235-0176

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 700 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

; ORIGINAL SOURCE:

; ORGANISM: Plasmodium falciparum

; SEQUENCE DESCRIPTION: SEQ ID NO: 10:

US-10-153-273-10

Query Match 5.0%; Score 325; DB 2; Length 700;

Best Local Similarity 24.4%; Pred. No. 5.8e-14;

Matches 120; Conservative 70; Mismatches 161; Indels 140; Gaps 22;

Qy 478 GTCEPERRRLTLC-----LGRTVLLHRGHEEDYKEHLLGASIIYEAQLLKYY-----KE 525

Db 8 GACAPYRRLHLCDYNLESIDTSTTHKLLE-----VCAAKYEGNSINTHYTHQHTNE 62

Qy 526 KBNALCSIIONSADLADIIGSDIISKDY-----YGKMEENL-----NKVAKDKKRNE 575

Db 63 DSASQLCTVLARSFADIGDIVRGKDLVGYDNKEKEQRKLEQKLDIPKIHKDVMTN 122

Qy 576 ESLKIF-----REKWDENKENVWK-----VMSAVLNKNET----- 606

Db 123 GAQERYIDDAKGGDFFQLREDMWTNSRETVMKALICHAPKEANYFIKTACNVGKTNGQC 182

Qy 607 -CKDYD---KFQKIPOFLRFKWEKGGDFCEKREKIYSPESPKECKK-----K 651

Db 183 HCIGGDVPTYFDYVFOYLRWFEEAEDFCRKKKKL---ENLQKQRDYEQNLYCSNGY 239

Qy 652 DCDE-----NTCKNKCSEYKKWIDLKSEYKQVDKYTKD-----KNK 689

Db 240 DCTKIYKKGKLVIGHECTNCVWCRMYETWIDNQKEFLKQKRKYETEISGGSGKSPK 299

Qy 690 KMYDNIDEVKNKEANVYLKESKECDVNFDD-----KIFNESPNYEYEDMC-----K 736

Db 300 RTKRAARSSSSDDNGYESKFYKLEKGVQDVDFKFLKLNK-----EGICQKQPQVGN 354

Qy 737 KCDEIKYLNIEIKYKPT--KHDIYDIDTFSDTFGDGTPIISANINAEQOQSGKDTSTNGSE 794

Db 355 KADNVDTNE-KYVKTFSTCEPCWGLEKGGPPWKV-----KGDKTCSAKTK 405

Qy 795 T-----SDSPVSHPESDAAINVEKLSGDESSSETRGIL-----DINDPVSVTNNVNE 841

Db 406 TYDPKNITDIPVLPDKSQNNILKYNFCCKGAPGGQIKKWQCYDEHRPSSKNNNC 465

Qy 842 VHDASN--TQG 850

Db 466 VEGTWKFTQG 476

RESULT 35

US-10-087-013-10

; Sequence 10, Application US/10087013

; Patent No. 6855323

; GENERAL INFORMATION:

; APPLICANT: Arthur Scherf

; APPLICANT: Louis H. Miller

; APPLICANT: Benoit Gamain

; APPLICANT: Dror I. Baruch

; APPLICANT: Pierre Buffet

; APPLICANT: Christine Scheidig

; APPLICANT: Jurg Gysin

; APPLICANT: Bruno Pouvelle

; APPLICANT: No. 6855323utaka Fujii

; APPLICANT: Joseph Smith

; TITLE OF INVENTION: IDENTIFICATION OF THE DOMAIN OF

; PLASMODIUM FALCIPARUM ERYTHROCYTE MEMBRANE PROTEIN 1

; (PEMP1) THAT MEDIATES ADHESION TO CHONDROITIN SULFATE A

; FILE REFERENCE: NIH176.001C1

; CURRENT APPLICATION NUMBER: US/10/087,013

; CURRENT FILING DATE: 2002-02-21

; PRIOR APPLICATION NUMBER: PCT/US00/24195

; PRIOR FILING DATE: 2000-09-01

; PRIOR APPLICATION NUMBER: 60/152,023

; PRIOR FILING DATE: 1999-09-01

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 10

; TYPE: PRT

; LENGTH: 311

; ORGANISM: Plasmodium falciparum

US-10-087-013-10

RESULT 39

US-08-487-826B-27
; Sequence 27, Application US/08487826B
; Patent No. 5993827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,826B
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001Cp1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 277 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
US-08-487-826B-27

Query Match 4.7%; Score 301.5; DB 1; Length 277;
Best Local Similarity 31.1%; Pred. No. 6.5e-13;
Matches 64; Conservative 25; Mismatches 112; Indels 5; Gaps 3;

Qy 468 CTGTFTNKFPGTCPPRRQTLCLGRITLLHRGHEEDYKEHLGLGASIIYEAQLLKYYKKEKD 527
Db 15 CXXXXXXXVCPVPPRRQELCLGNIXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 74

Qy 528 ENALCSIIQNSYADLADIIKGSIIIDYKYGKMEENLNKVNKKRNEESLKIFREKWD 587
Db 75 XXEYCKIINTFADIRDIIGTDTYNDLSNRXXXXXXXXXXXXXXXXXXXXKNDKLFDEWVK 134

Qy 588 ENKENVWKMVASLKNKCTCKDYDFQKIPOFLRWPKWGDGDDFCCKEKEKIKYSPESFKVE 647
Db 135 VIKDWNVNIS-WFXXXXXC-XXXXXXXIPOFFRWFSEWGDYCODTKMI---ETLKVE 189

Qy 648 CKKXDCDENTCKNCKSEYKKWIDLKK 673
Db 190 CXXXXCXDDNCKSKNSYKEWISKKK 215

RESULT 40

US-09-210-288-15
; Sequence 15, Application US/09210288
; Patent No. 6392026
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/210,288
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fuller, Michael
; REGISTRATION NUMBER: 36,516
; REFERENCE/DOCKET NUMBER: NIH121.1FWDV1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 277 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
US-09-210-288-15

Query Match 4.7%; Score 301.5; DB 2; Length 277;
Best Local Similarity 31.1%; Pred. No. 6.5e-13;
Matches 64; Conservative 25; Mismatches 112; Indels 5; Gaps 3;

Qy 468 CTGTFTNKFPGTCPPRRQTLCLGRITLLHRGHEEDYKEHLGLGASIIYEAQLLKYYKKEKD 527
Db 15 CXXXXXXXVCPVPPRRQELCLGNIXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 74

Qy 528 ENALCSIIQNSYADLADIIKGSIIIDYKYGKMEENLNKVNKKRNEESLKIFREKWD 587
Db 75 XXEYCKIINTFADIRDIIGTDTYNDLSNRXXXXXXXXXXXXXXXXXXXXKNDKLFDEWVK 134

Qy 588 ENKENVWKMVASLKNKCTCKDYDFQKIPOFLRWPKWGDGDDFCCKEKEKIKYSPESFKVE 647
Db 135 VIKDWNVNIS-WFXXXXXC-XXXXXXXIPOFFRWFSEWGDYCODTKMI---ETLKVE 189

Qy 648 CKKXDCDENTCKNCKSEYKKWIDLKK 673
Db 190 CXXXXCXDDNCKSKNSYKEWISKKK 215

Search completed: November 21, 2005, 20:49:10
Job time : 75 secs

This Page Blank (uspto)

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2005, 20:26:58 ; Search time 257 Seconds
(without alignment)
3321.748 Million cell updates/sec

Title: US-10-677-980-2
Perfect score: 6481
Sequence: 1 MKGYFNIFLIPLIFLYNVI.....VQETNISDYSEYNNKMY 1210

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 2166443 seqs, 705528306 residues
Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6481	100.0	1210	2	Q96724 PLAF7
2	6471	99.8	1210	2	Q8WS31 PLAF7
3	6471	99.8	1210	2	Q76NM5 PLAF7
4	6091	94.0	1137	2	Q8WM32 PLAF7
5	6085	93.9	1137	2	Q8MT19 PLAF7
6	6080	93.8	1137	2	Q8MT18 PLAF7
7	6077	93.8	1137	2	Q8MT16 PLAF7
8	6077	93.8	1137	2	Q8MT17 PLAF7
9	5606.5	86.5	1282	2	C5SEK8 PLARE
10	3374	52.1	616	2	Q8MM04 PLAF7
11	3368	52.0	616	2	Q8MM45 PLAF7
12	3367	52.0	616	2	Q8T918 PLAF7
13	3365	51.9	616	2	Q8I049 PLAF7
14	3364	51.9	616	2	Q8MM74 PLAF7
15	3364	51.9	616	2	Q8T9N3 PLAF7
16	3363	51.9	616	2	Q8I787 PLAF7
17	3360	51.8	616	2	Q8I788 PLAF7
18	3359	51.8	616	2	Q8I786 PLAF7
19	3319	51.2	606	2	Q66PM7 PLAF7
20	3309	51.1	606	2	Q66PL8 PLAF7
21	3144	48.5	616	2	Q8I760 PLARE
22	1700	26.2	1475	2	Q25842 PLAF7
23	1596.5	24.6	1462	2	Q8IB88 PLAF7
24	1557	24.0	1435	1	EBAL PLAF7
25	1547	23.9	1435	2	Q9NG63 PLAF7
26	1537.5	23.7	1421	2	Q05644 PLAF7
27	1524	23.5	1402	2	Q9SV11 PLAF7
28	1514.5	23.4	1433	2	Q9N9G9 PLARE
29	1210.5	18.7	614	2	Q8I796 PLAF7
30	1209.5	18.7	616	2	Q8I791 PLAF7
31	1209.5	18.7	616	2	Q8I797 PLAF7

32	1207.5	18.6	614	2	Q8I7A0 PLAF7	Q8I7A0 plasmodium
33	1206.5	18.6	616	2	P90585 PLAF7	P90585 plasmodium
34	1206.5	18.6	616	2	Q25739 PLAF7	Q25739 plasmodium
35	1204.5	18.6	614	2	Q25736 PLAF7	Q25736 plasmodium
36	1204.5	18.6	614	2	Q8I790 PLAF7	Q8I790 plasmodium
37	1204.5	18.6	616	2	Q8I0C8 PLAF7	Q8I0C8 plasmodium
38	1204.5	18.6	616	2	Q27390 PLAF7	Q27390 plasmodium
39	1203.5	18.6	616	2	Q27248 PLAF7	Q27248 plasmodium
40	1202.5	18.6	616	2	Q27361 PLAF7	Q27361 plasmodium
41	1202.5	18.6	616	2	Q8I795 PLAF7	Q8I795 plasmodium
42	1199.5	18.5	616	2	Q25735 PLAF7	Q25735 plasmodium
43	1198.5	18.5	614	2	Q8I0U2 PLAF7	Q8I0U2 plasmodium
44	1198.5	18.5	616	2	Q8I792 PLAF7	Q8I792 plasmodium
45	1198.5	18.5	616	2	Q8I793 PLAF7	Q8I793 plasmodium

ALIGNMENTS

RESULT 1
Q96724 PLAF7
ID Q96724 PLAF7 PRELIMINARY; PRT; 1210 AA.
AC Q96724;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Erythrocyte binding protein BAEBL.
GN Name=baebl;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Indochina III/CDC;
RX MEDLINE=21211106; PubMed=11309486; DOI=10.1073/pnas.081075398;
RA Mayer D.C., Kaneko O., Hudson-Taylor D.E., Reid M.E., Miller L.H.;
RT "Characterization of a Plasmodium falciparum erythrocyte-binding
protein paralogous to EBA-175.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:5222-5227(2001).
DR EMBL; AF332918; AAK49521.1; -, Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; P:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy binding.
DR Pfam; PF05424; Duffy binding, 1.
SQ SEQUENCE 1210 AA; 140527 MW; F0E56B397C8E8271 CRC64;

Query Match	100.0%;	Score	6481;	DB 2;	Length	1210;
Best Local Similarity	100.0%;	Pred. No.	6.1e-264;			
Matches 1210;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;					
Qy	1	MKGYNFIYFLIPLIFLYNVRINESIIGRTLYNRQDESSDISRVNSPELNHNHNTIYDS	60			
Db	1	MKGYNFIYFLIPLIFLYNVRINESIIGRTLYNRQDESSDISRVNSPELNHNHNTIYDS	60			
Qy	61	DYEDVNNKLSNFVENKSVKKKSLSPINNKTYSYDIIPPSYSYRNDKFNLSNENSG	120			
Db	61	DYEDVNNKLSNFVENKSVKKKSLSPINNKTYSYDIIPPSYSYRNDKFNLSNENSG	120			
Qy	121	NTNSNFPANTSEISIGKDNKQYTFIQKTHLFAAGIKRKSIIKWTICRENSKITYCVDPDRK	180			
Db	121	NTNSNFPANTSEISIGKDNKQYTFIQKTHLFAAGIKRKSIIKWTICRENSKITYCVDPDRK	180			
Qy	181	IQLCVANFLNRLSRLETWEKFEIILSVNTEAKLLYNKQEKDPSIFCNELRNSFSPRSS	240			
Db	181	IQLCVANFLNRLSRLETWEKFEIILSVNTEAKLLYNKQEKDPSIFCNELRNSFSPRSS	240			
Qy	241	FIGDDMDFGNTDRVKGYINTKESDYKKNVKNLNNIKKEWKEKNKANLWNHMI VNHKG	300			
Db	241	FIGDDMDFGNTDRVKGYINTKESDYKKNVKNLNNIKKEWKEKNKANLWNHMI VNHKG	300			
Qy	301	NISKECAIIPAEPPQINLWIKWENENFMEKRLFLNIDKDCVENKKEACFGGCRLPSCS	360			
Db	301	NISKECAIIPAEPPQINLWIKWENENFMEKRLFLNIDKDCVENKKEACFGGCRLPSCS	360			

Db 301 NISKECAIIPAEBPQINLIWKWENFLMEKRLFLNIKDKCVENKYYEACFGCRLPCS 360
Qy 361 SYTSFMKSKTQMEVLNLYKKQSGVDKNNFLNDLFKKNKNDLDDFFKNEKEYDDLCD 420
Db 361 SYTSFMKSKTQMEVLNLYKKQSGVDKNNFLNDLFKKNKNDLDDFFKNEKEYDDLCD 420
Qy 421 CRYTATIIKSLFNGPAKNDVDIASQINVDLRGFCGNYKSNKSNKWNCTGTFTNKPFGTC 480
Db 421 CRYTATIIKSLFNGPAKNDVDIASQINVDLRGFCGNYKSNKSNKWNCTGTFTNKPFGTC 480
Qy 481 EPPRROTCLGRYLLHRGHEEDYKEHLLGASIEYAQLLYKKEKDNALCSIIQNSYA 540
Db 481 EPPRROTCLGRYLLHRGHEEDYKEHLLGASIEYAQLLYKKEKDNALCSIIQNSYA 540
Qy 541 DLADIIKGSDDIIKDYGGKMEENLVNKKKNEESLKIIFREKWDNKENYKVMASV 600
Db 541 DLADIIKGSDDIIKDYGGKMEENLVNKKKNEESLKIIFREKWDNKENYKVMASV 600
Qy 601 LKNKETCKDYDFOKIPIQFLRWFKWGGDDFCEKKEKIYFSEPKVECKKDCDENTCKN 660
Db 601 LKNKETCKDYDFOKIPIQFLRWFKWGGDDFCEKKEKIYFSEPKVECKKDCDENTCKN 660
Qy 661 KCSEYKKWIDLKSEYKQVDTYKDKNKKMYDNIDEVKKEANVYLKESKECKDVNFD 720
Db 661 KCSEYKKWIDLKSEYKQVDTYKDKNKKMYDNIDEVKKEANVYLKESKECKDVNFD 720
Qy 721 DKIFNESPNEYEDMCKCDEIKYLNIEIKYPTKHDYIDIDTFSDTFDGGTPIISINANINE 780
Db 721 DKIFNESPNEYEDMCKCDEIKYLNIEIKYPTKHDYIDIDTFSDTFDGGTPIISINANINE 780
Qy 781 QOSGKDTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSSETRGLDINDPSVTNNV 840
Db 781 QOSGKDTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSSETRGLDINDPSVTNNV 840
Qy 841 EVHDSATQGSVNTSDITNGHSSSLNRTTNAQDIKIGSGNQSDNQNSHSSDNG 900
Db 841 EVHDSATQGSVNTSDITNGHSSSLNRTTNAQDIKIGSGNQSDNQNSHSSDNG 900
Qy 901 SLTIGQVPSDNTQNTYDSQNPHRDTPNALASPSDDKINEIEGFSRSDSENGRGTTS 960
Db 901 SLTIGQVPSDNTQNTYDSQNPHRDTPNALASPSDDKINEIEGFSRSDSENGRGTTS 960
Qy 961 NTHDVRNTIVSERVNSHDFIRGMANNAAHQYITQIENNGIIRQOESAGSNVYKD 1020
Db 961 NTHDVRNTIVSERVNSHDFIRGMANNAAHQYITQIENNGIIRQOESAGSNVYKD 1020
Qy 1021 NPKESNFSSENHKKNIQEVNSRDTKVRBEIILSKONKCNNEYSMEYCTYSDERNSSP 1080
Db 1021 NPKESNFSSENHKKNIQEVNSRDTKVRBEIILSKONKCNNEYSMEYCTYSDERNSSP 1080
Qy 1081 GPCSRERKKLCCQISDYCLKFNFYFIEYNYCIKSEIKSPYKCFKSEGOSSIPYPAAG 1140
Db 1081 GPCSRERKKLCCQISDYCLKFNFYFIEYNYCIKSEIKSPYKCFKSEGOSSIPYPAAG 1140
Qy 1141 GILVIVILLSSARMKSENEVDIGESNTEATPEENNYLNKLSRIENQVQETNISDYS 1200
Db 1141 GILVIVILLSSARMKSENEVDIGESNTEATPEENNYLNKLSRIENQVQETNISDYS 1200
Qy 1201 EYNYNEKNMY 1210
Db 1201 EYNYNEKNMY 1210

RESULT 2

Q8WS31.PLAFA
ID Q8WS31.PLAFA PRELIMINARY; PRT; 1210 AA.
AC Q8WS31;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DE Erythrocyte binding antigen 140.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3D7;
RA Thompson J.K., Triglia T., Reed M.B., Cowman A.F.;
RT "A novel ligand from Plasmodium falciparum that binds to a sialic
RT acid-containing receptor on the surface of human erythrocytes.";
RL Mol. Microbiol. 0:0-0(2002).
DR EMBL; AF384554; AAL58319.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:lipathogenesis; IEA.
DR InterPro; IPR008602; Duff binding; IEA.
DR Pfam; PF05424; Duff binding; 1.
SQ SEQUENCE 1210 AA; 140595 MW; PF44FBB8DA69D605 CRC64;

Query Match 99.8%; Score 6471; DB 2; Length 1210;

Best Local Similarity 99.8%; Pred. No. 1.6e-263;

Matches 1207; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MKGYFNIYFLIPLIFLYNVIRINESIIGRTLYNRQDESSDISRVNSPELNNHKTNIYDS 60

Db 1 MKGYFNIYFLIPLIFLYNVIRINESIIGRTLYNRQDESSDISRVNSPELNNHKTNIYDS 60

Qy 61 DYEDVNNKLINSFVENKSVKKGSLSFNNKTKSYDIIIPPSYRYNDKFNLSLENEDNSG 120

Db 61 DYEDVNNKLINSFVENKSVKKGSLSFNNKTKSYDIIIPPSYRYNDKFNLSLENEDNSG 120

Qy 121 NTSNNFANTSEISIGKDNKQYTFIOKRTHLFACGIRKRSIKWICRENSSEKITVCPVDRK 180

Db 121 NTSNNFANTSEISIGKDNKQYTFIOKRTHLFACGIRKRSIKWICRENSSEKITVCPVDRK 180

Qy 181 IQLCVANFLNSRLTMEKFEFLISVNTAEKLLYNKNEGKOPSIFCNELRNSFSFPRS 240

Db 181 IQLCVANFLNSRLTMEKFEFLISVNTAEKLLYNKNEGKOPSIFCNELRNSFSFPRS 240

Qy 241 FIGDDMDFGNTDRVKGYNTRKPSDYKQVKEVLEKLANIKKEWKEKCANLWHEMIVNHKG 300

Db 241 FIGDDMDFGNTDRVKGYNTRKPSDYKQVKEVLEKLANIKKEWKEKCANLWHEMIVNHKG 300

Qy 301 NISKECAIIPAEBPQINLIWKWENFLMEKRLFLNIKDKCVENKYYEACFGCRLPCS 360

Db 301 NISKECAIIPAEBPQINLIWKWENFLMEKRLFLNIKDKCVENKYYEACFGCRLPCS 360

Qy 361 SYTSFMKSKTQMEVLNLYKKQSGVDKNNFLNDLFKKNKNDLDDFFKNEKEYDDLCD 420

Db 361 SYTSFMKSKTQMEVLNLYKKQSGVDKNNFLNDLFKKNKNDLDDFFKNEKEYDDLCD 420

Qy 421 CRYTATIIKSLFNGPAKNDVDIASQINVDLRGFCGNYKSNKSNKWNCTGTFTNKPFGTC 480

Db 421 CRYTATIIKSLFNGPAKNDVDIASQINVDLRGFCGNYKSNKSNKWNCTGTFTNKPFGTC 480

Qy 481 EPPRROTCLGRYLLHRGHEEDYKEHLLGASIEYAQLLYKKEKDNALCSIIQNSYA 540

Db 481 EPPRROTCLGRYLLHRGHEEDYKEHLLGASIEYAQLLYKKEKDNALCSIIQNSYA 540

Qy 541 DLADIIKGSDDIIKDYGGKMEENLVNKKKNEESLKIIFREKWDNKENYKVMASV 600

Db 541 DLADIIKGSDDIIKDYGGKMEENLVNKKKNEESLKIIFREKWDNKENYKVMASV 600

Qy 601 LKNKETCKDYDFOKIPIQFLRWFKWGGDDFCEKKEKIYFSEPKVECKKDCDENTCKN 660

Db 601 LKNKETCKDYDFOKIPIQFLRWFKWGGDDFCEKKEKIYFSEPKVECKKDCDENTCKN 660

Qy 661 KCSEYKKWIDLKSEYKQVDTYKDKNKKMYDNIDEVKKEANVYLKESKECKDVNFD 720

Db 661 KCSEYKKWIDLKSEYKQVDTYKDKNKKMYDNIDEVKKEANVYLKESKECKDVNFD 720

Qy 721 DKIFNESPNEYEDMCKCDEIKYLNIEIKYPTKHDYIDIDTFSDTFDGGTPIISINANINE 780

Db 721 DKIFNESPNEYEDMCKCDEIKYLNIEIKYPTKHDYIDIDTFSDTFDGGTPIISINANINE 780

Qy 781 QOSGKDTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSSETRGLDINDPSVTNNV 840

```
Db 781 QSGKDTNTGNSSETSDSPVSHPEPESDAAINVEKLSGDESSETRGILDINDPSVTNNVN 840
Qy 841 EVHDASNTQGSVNTSDITNGHSSSLNRTTNAODIKIGRSGNQSDNOENSSHSSDMSG 900
Db 841 EVHDASNTQGSVNTSDITNGHSSSLNRTTNAODIKIGRSGNQSDNOENSSHSSDMSG 900
Qy 901 SLTIGQVPSDNTQNTYDSQNPDRDTPNALASLPSDDKINEIEGFDSSRDSSENGRDTTS 960
Db 901 SLTIGQVPSDNTQNTYDSQNPDRDTPNALASLPSDDKINEIEGFDSSRDSSENGRDTTS 960
Qy 961 NTHDVRRTNIVSERVNSHDFIRNGMANNNAHQYITQIENNGIIRGOESAGNSVNYKD 1020
Db 961 NTHDVRRTNIVSERVNSHDFIRNGMANNNAHQYITQIENNGIIRGOESAGNSVNYKD 1020
Qy 1021 NPKRSNFSSNDHKNIQYNSRDTKRVREBIIKLSKQNKCNNEYSMEYCTYSDERNSSP 1080
Db 1021 NPKRSNFSSNDHKNIQYNSRDTKRVREBIIKLSKQNKCNNEYSMEYCTYSDERNSSP 1080
Qy 1081 GPCSRERKKLCCQISDYCLAYFNFYISIEYNYNCIKSEIKSPYKCFKSEGGSSIPYFAAG 1140
Db 1081 GPCSRERKKLCCQISDYCLAYFNFYISIEYNYNCIKSEIKSPYKCFKSEGGSSIPYFAAG 1140
Qy 1141 GILVIVILLSSASRMGKSNEEYDIGESNIEATFEENNYLNKLSRIFNQVQETNISDYS 1200
Db 1141 GILVIVILLSSASRMGKSNEEYDIGESNIEATFEENNYLNKLSRIFNQVQETNISDYS 1200
Qy 1201 EYNTNEKMY 1210
Db 1201 EYNTNEKMY 1210

RESULT 3
ID Q76NMS_PLAF7 PRELIMINARY; PRT; 1210 AA.
AC Q76NMS;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Erythrocyte binding antigen 140.
GN Name=WAL13P1.60;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3D7;
RA Harris B.; Lennard N.; Clark L.; Line A.; Barron A.; Corton C.;
RA Berriman M.; Pain A.; Hall N.; Atkin R.; Chillingworth C.; Doggett J.;
RA Ormond D.; Sanders M.; Hayes R.; Hall S.; Quail M.; Barrall B.;
RL Submitted (SRP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL844509; CAD52266.1; -; Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy binding.
DR Pfam; PF05424; Duffy binding; 1.
SQ SEQUENCE 1210 AA; 140595 MW; FF44FB8DA69D605 CRC64;

Query Match 99.8%; Score 6471; DB 2; Length 1210;
Best Local Similarity 99.8%; Pred. No. 1.6e-263;
Matches 1207; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MKGFNYFLPLFLPLNYNIRINESIGRTLYNQDESSDLSRVNSPELNHNHTNIYDS 60
Db 1 MKGFNYFLPLFLPLNYNIRINESIGRTLYNQDESSDLSRVNSPELNHNHTNIYDS 60
Qy 61 DYEDVNNKLSFVFNKSVKKRSLSFNNKTKSDIIPPSYVRNDKFNLSNEONSG 120
Db 61 DYEDVNNKLSFVFNKSVKKRSLSFNNKTKSDIIPPSYVRNDKFNLSNEONSG 120
Qy 121 NTNSNNFANTSEISIGKDNKQYTFIQKRTLFACGIKRKSIIKWICRENSEKITVCVPRK 180
Db 121 NTNSNNFANTSEISIGKDNKQYTFIQKRTLFACGIKRKSIIKWICRENSEKITVCVPRK 180
```

```
Db 121 NTNSNNFANTSEISIGKDNKQYTFIQKRTLFACGIKRKSIIKWICRENSEKITVCVPRK 180
Qy 181 IQLCVANFLSRLTMEKFEIPLISVNTAEKLYNKNEGKDPISFCNELANSFSDPRESS 240
Db 181 IQLCVANFLSRLTMEKFEIPLISVNTAEKLYNKNEGKDPISFCNELANSFSDPRESS 240
Qy 241 FIGDMDFGGNTDRVKGYINTKFSDDYKKNVEKLANNIKKEWEKKNKANLWNHMI VNHKG 300
Db 241 FIGDMDFGGNTDRVKGYINKKFSDDYKKNVEKLANNIKKEWEKKNKANLWNHMI VNHKG 300
Qy 301 NISKECAIIIPAEERQINLWIKWENENFLMEKRLPLANKDKCVENKKEVYACFGCRLPCS 360
Db 301 NISKECAIIIPAEERQINLWIKWENENFLMEKRLPLANKDKCVENKKEVYACFGCRLPCS 360
Qy 361 SYTSFMKKSQTOMEVLNLYKKNKSGVDKNNFLNDLPKNNKNDLDDPFFKNEKEYDLDLDCD 420
Db 361 SYTSFMKKSQTOMEVLNLYKKNKSGVDKNNFLNDLPKNNKNDLDDPFFKNEKEYDLDLDCD 420
Qy 421 CRYTATIIKSGFLNGPAKNDVDIASQINVDLARGFCGNYKSNNEKSWNCTGTTNKFPGTC 480
Db 421 CRYTATIIKSGFLNGPAKNDVDIASQINVDLARGFCGNYKSNNEKSWNCTGTTNKFPGTC 480
Qy 481 EPPRQTLCLGRTYLLHRGHEEDYKEHLGASIIYEAQLLYKKEKEDENALCSIIQNSYA 540
Db 481 EPPRQTLCLGRTYLLHRGHEEDYKEHLGASIIYEAQLLYKKEKEDENALCSIIQNSYA 540
Qy 541 DLADIIGSDIIKDYGYKKMEENLNKYNKDKKRNEESLKI PREKQWDEKENVKMWMSAV 600
Db 541 DLADIIGSDIIKDYGYKKMEENLNKYNKDKKRNEESLKI PREKQWDEKENVKMWMSAV 600
Qy 601 LKNETCKDYDKFOKIPQFLRWPKWGGDDFCEKKEKIIYSPESPKVECKKDCDENTCKN 660
Db 601 LKNETCKDYDKFOKIPQFLRWPKWGGDDFCEKKEKIIYSPESPKVECKKDCDENTCKN 660
Qy 661 KCSEYKWIIDLKSEYEKQVDKYTKDKNKQWYDNIDEVKNKEANVYLKEKSECKQVDFD 720
Db 661 KCSEYKWIIDLKSEYEKQVDKYTKDKNKQWYDNIDEVKNKEANVYLKEKSECKQVDFD 720
Qy 721 DKIFNESPNEVEDMCKKDEIKYLINEIKYPTKHDIIYDIDTFSDTFGDTGPISINANINE 780
Db 721 DKIFNESPNEVEDMCKKDEIKYLINEIKYPTKHDIIYDIDTFSDTFGDTGPISINANINE 780
Qy 781 QSGKDTNTGNSSETSDSPVSHPEPESDAAINVEKLSGDESSETRGILDINDPSVTNNVN 840
Db 781 QSGKDTNTGNSSETSDSPVSHPEPESDAAINVEKLSGDESSETRGILDINDPSVTNNVN 840
Qy 841 EVHDASNTQGSVNTSDITNGHSSSLNRTTNAODIKIGRSGNQSDNOENSSHSSDMSG 900
Db 841 EVHDASNTQGSVNTSDITNGHSSSLNRTTNAODIKIGRSGNQSDNOENSSHSSDMSG 900
Qy 901 SLTIGQVPSDNTQNTYDSQNPDRDTPNALASLPSDDKINEIEGFDSSRDSSENGRDTTS 960
Db 901 SLTIGQVPSDNTQNTYDSQNPDRDTPNALASLPSDDKINEIEGFDSSRDSSENGRDTTS 960
Qy 961 NTHDVRRTNIVSERVNSHDFIRNGMANNNAHQYITQIENNGIIRGOESAGNSVNYKD 1020
Db 961 NTHDVRRTNIVSERVNSHDFIRNGMANNNAHQYITQIENNGIIRGOESAGNSVNYKD 1020
Qy 1021 NPKRSNFSSNDHKNIQYNSRDTKRVREBIIKLSKQNKCNNEYSMEYCTYSDERNSSP 1080
Db 1021 NPKRSNFSSNDHKNIQYNSRDTKRVREBIIKLSKQNKCNNEYSMEYCTYSDERNSSP 1080
Qy 1081 GPCSRERKKLCCQISDYCLAYFNFYISIEYNYNCIKSEIKSPYKCFKSEGGSSIPYFAAG 1140
Db 1081 GPCSRERKKLCCQISDYCLAYFNFYISIEYNYNCIKSEIKSPYKCFKSEGGSSIPYFAAG 1140
Qy 1141 GILVIVILLSSASRMGKSNEEYDIGESNIEATFEENNYLNKLSRIFNQVQETNISDYS 1200
Db 1141 GILVIVILLSSASRMGKSNEEYDIGESNIEATFEENNYLNKLSRIFNQVQETNISDYS 1200
Qy 1201 EYNTNEKMY 1210
Db 1201 EYNTNEKMY 1210
```

Db	601	LKNKETCKDYDKFKIPIQFLRWFKEGGDDFCERKEKIYSPFSFKVECKKDCDENTCKN	660
Qy	661	KCSEYKKWIDLKKSSEYKQVDKYTKDNKKQMYDNIDEVKNKANYLVLEKSKCECKDVNFD	720
Db	661	KCSEYKKWIDLKKSSEYKQVDKYTKDNKKQMYDNIDEVKNKANYLVLEKSKCECKDVNFD	720
Qy	721	DKIFNESPNEYEDMCKKCDKDEIKYLNEIKYPTKDKIYDIDTFSDFGDTGISINANINE	780
Db	721	DKIFNESPNEYEDMCKKCDKDEIKYLNEIKYPTKDKIYDIDTFSDFGDTGISINANINE	780
Qy	781	QOSGKDTSTNTGNSSETSDSPVSHPEPSDAAINVEKLSGDESSSETRGILDINDPSVTNNVN	840
Db	781	QOSGKDTSTNTGNSSETSDSPVSHPEPSDAAINVEKLSGDESSSETRGILDINDPSVTNNVN	840
Qy	841	EYHDASNTQGSVNSNTSDITNGHSESSLNRTTNAQDIKIGRSNEQSDNQENSHSSDMSG	900
Db	841	EYHDASNTQGSVNSNTSDITNGHSESSLNRTTNAQDIKIGRSNEQSDNQENSHSSDMSG	900
Qy	901	SLTIQGVPESDNTQNTYDSQNPHRDTPNALASLPDDDKINELEGDSSDSNGRGDTTS	960
Db	901	SLTIQGVPESDNTQNTYDSQNPHRDTPNALASLPDDDKINELEGDSSDSNGRGDTTS	960
Qy	961	NTHDVRRTNIVSERRVNSHDFIRNGMANNNAHQQYITQIENNGIIRGOESAGNSVNYKD	1020
Db	961	NTHDVRRTNIVSERRVNSHDFIRNGMANNNAHQQYITQIENNGIIRGOESAGNSVNYKD	1020
Qy	1021	NPKRSFSSENDHKKNIQEYNSRDTKRVREELIILSKONKCNNEYSMEYCTTSDERNSSP	1080
Db	1021	NPKRSFSSENDHKKNIQEYNSRDTKRVREELIILSKONKCNNEYSMEYCTTSDERNSSP	1080
Qy	1081	GPCSEERKKLCCQISDYCLKYFNFSIYNYNCIKSEIKSPYKCFKSEGGQSSI	1134
Db	1081	GPCSEERKKLCCQISDYCLKYFNFSIYNYNCIKSEIKSPYKCFKSEGGQSSM	1134
RESULT 5			
ID	Q8MT19	PLAFA PRELIMINARY;	PRT; 1137 AA.
AC	Q8MT19		
DT	01-OCT-2002	(TrEMBLrel. 22, Created)	
DT	01-OCT-2002	(TrEMBLrel. 22, Last sequence update)	
DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)	
DE	Erythrocyte-binding protein.		
OS	Plasmodium falciparum.		
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
OX	NCBI_TaxID=5833;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RA	Mayer D.G., Mu J.-B., Feng X., Su X.-Z., Miller L.H.;		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL;	AY099886;	AAM51622.1; -; Genomic DNA.
DR	GO;	GO:0016021;	C:integral to membrane; IEA.
DR	GO;	GO:0004872;	F:receptor activity; IEA.
DR	GO;	GO:0009405;	P:pathogenesis; IEA.
DR	InterPro; IPR008602; Duffy binding.		
DR	Pfam; PF05424; Duffy binding; 1.		
SQ	SEQUENCE 1137 AA; 132218 MW; 8E1DC6CFDDA6083 CRC64;		
Query Match 94.0%; Score 6091; DB 2; Length 1137;			
Best Local Similarity 99.9%; Pred. No. 1.3e-247;			
Matches 1133; Conservative 1; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MKGYNFYFLPLFLFLYNVIRINESII	GRTLYNRQDESSDISRVNSPELN
Db	1	MKGYNFYFLPLFLFLYNVIRINESII	GRTLYNRQDESSDISRVNSPELN
Qy	61	DYEDVNNKLINSFVENKSVKKRSLSP	INNKTKSYDIIPPSYSYRNDKFN
Db	61	DYEDVNNKLINSFVENKSVKKRSLSP	INNKTKSYDIIPPSYSYRNDKFN
Qy	121	NTNSNFPANTSEISIGKDNKQYTFIQ	RTHLFCAGIKRKSIIKWCIRENSE
Db	121	NTNSNFPANTSEISIGKDNKQYTFIQ	RTHLFCAGIKRKSIIKWCIRENSE
Qy	181	IQLCVANFLNSRLTMEKPKFELIISV	NTAKLLYNKNEGKPSIFCNELRNS
Db	181	IQLCVANFLNSRLTMEKPKFELIISV	NTAKLLYNKNEGKPSIFCNELRNS
Qy	241	FIGDDMDPGGNTDRVKGYNITKFS	DYKKEKKNVEKLANIKKEWKEK
Db	241	FIGDDMDPGGNTDRVKGYNITKFS	DYKKEKKNVEKLANIKKEWKEK
Qy	301	NISKECAIIIAEBEQIINLWKENNE	FLMEKRLFLNIKDKCVENKKEAC
Db	301	NISKECAIIIAEBEQIINLWKENNE	FLMEKRLFLNIKDKCVENKKEAC
Qy	361	SYTSPMKKSTQMEVLTNLYKKNSG	VDKXNFLNDLFPKNNKNDLDD
Db	361	SYTSPMKKSTQMEVLTNLYKKNSG	VDKXNFLNDLFPKNNKNDLDD
Qy	421	CRYTATIIKSLFNGPAKNDVDIASQ	INVDLGRFGCNYKNNKSNCTGT
Db	421	CRYTATIIKSLFNGPAKNDVDIASQ	INVDLGRFGCNYKNNKSNCTGT
Qy	481	EPFRQTLCLGRTYLLHRGHEEDYK	HLGASIIYEAQLLYKKEKDNAL
Db	481	EPFRQTLCLGRTYLLHRGHEEDYK	HLGASIIYEAQLLYKKEKDNAL
Qy	541	DLADIIKGSIIIDYDYGKQMEENL	KVYNKDKGRNEESLKIIFREKWD
Db	541	DLADIIKGSIIIDYDYGKQMEENL	KVYNKDKGRNEESLKIIFREKWD
Qy	601	LKNKETCKDYDKFKIPIQFLRWFKEGGDDFCERKEKIYSPFSFKVECKKDCDENTCKN	660

Db 121 NTNSNFPANTSEISIGKNDKQYTFIQKRTFLFACGIRKRSIKWICRENSSEKITVCPDRK 180
Qy 181 IQLCVANFLNRLTMEKFKEIFLISVNTAEKLLYNKNEGKPSIFCNELRNSFSDFRSS 240
Db 181 IQLCVANFLNRLTMEKFKEIFLISVNTAEKLLYNKNEGKPSIFCNELRNSFSDFRSS 240
Qy 241 FIGDDMPGGTDRVKGINTKFSYDYKQKVEKLNKIKKEWKEKKNANLWNIHVMHKG 300
Db 241 FIGDDMPGGTDRVKGINTKFSYDYKQKVEKLNKIKKEWKEKKNANLWNIHVMHKG 300
Qy 301 NISKECAIIPAEFPQINLWIKENENFLMEKRLFLNIDKCVENKKEKYEACFGCRLPCS 360
Db 301 NISKECAIIPAEFPQINLWIKENENFLMEKRLFLNIDKCVENKKEKYEACFGCRLPCS 360
Qy 361 SYTSFMKSKTQMEVLNLYKKQNSGVDKNNFLNDFKNNKNDLDDFFKNEKEYDLDL 420
Db 361 SYTSFMKSKTQMEVLNLYKKQNSGVDKNNFLNDFKNNKNDLDDFFKNEKEYDLDL 420
Qy 421 CRYTATIIKSLFNGPAKNDVDIASQINVDNLRGFGCNYSKNEKSWNCTGTFNKPFGTC 480
Db 421 CRYTATIIKSLFNGPAKNDVDIASQINVDNLRGFGCNYSKNEKSWNCTGTFNKPFGTC 480
Qy 481 EPPRQTLCLGRYTLHRRGHEEDYKEHLLGASIIYEAQLLKYKKEKDNALCSIIQNSYA 540
Db 481 EPPRQTLCLGRYTLHRRGHEEDYKEHLLGASIIYEAQLLKYKKEKDNALCSIIQNSYA 540
Qy 541 DLADIIKGSDDIIKDYKQKMEENLKNVNDKKNRNEESLKIPEKRWMDENKENVKWSAV 600
Db 541 DLADIIKGSDDIIKDYKQKMEENLKNVNDKKNRNEESLKIPEKRWMDENKENVKWSAV 600
Qy 601 LKNEKTCQYDKFQKIPOFLWFKWGGDDFCERKEKIIYSPESFKVECKKCDENTCKN 660
Db 601 LKNEKTCQYDKFQKIPOFLWFKWGGDDFCERKEKIIYSPESFKVECKKCDENTCKN 660
Qy 661 KCSYKWKWIDLKSEYKQVDTYKDKNKKQYNDIDEVONKEANVYLKESKECKQVNF 720
Db 661 KCSYKWKWIDLKSEYKQVDTYKDKNKKQYNDIDEVONKEANVYLKESKECKQVNF 720
Qy 721 DKIFNESPNEVEDMCKKDEIKYLNEIKYPTKHDIIYDIDTFSDTFGDTGPISINANINE 780
Db 721 DKIFNESPNEVEDMCKKDEIKYLNEIKYPTKHDIIYDIDTFSDTFGDTGPISINANINE 780
Qy 781 QQSGKDTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSSETRGILDINDPSVTNNV 840
Db 781 QQSGKDTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSSETRGILDINDPSVTNNV 840
Qy 841 EVHDSNTQGSVNTSDITNGHSSSLNRTTNAQDIKIGRSGNEQSDNQNSHSSDNG 900
Db 841 EVHDSNTQGSVNTSDITNGHSSSLNRTTNAQDIKIGRSGNEQSDNQNSHSSDNG 900
Qy 901 SLTIGQVPSDNTQNTYDSQNPDRDTPNALASLPSDDKINEIEGFDSSRDSSENGRDTTS 960
Db 901 SLTIGQVPSDNTQNTYDSQNPDRDTPNALASLPSDDKINEIEGFDSSRDSSENGRDTTS 960
Qy 961 NTHDVRRTNIVSERVNSHPDRNGMANNNAHQYITQIENNGIIRGOESAGNSVNYKD 1020
Db 961 NTHDVRRTNIVSERVNSHPDRNGMANNNAHQYITQIENNGIIRGOESAGNSVNYKD 1020
Qy 1021 NPKSNFSSNDHKKNIQEVNSRDTKRVREBIIKLSKQNKCNNSYMEYCTYSDERNSSP 1080
Db 1021 NPKSNFSSNDHKKNIQEVNSRDTKRVREBIIKLSKQNKCNNSYMEYCTYSDERNSSP 1080
Qy 1081 GPCSRERKKLCCQISDYCLKFNFPYSLEYVNCIKSBIKSPYKCFKSEGOSSI 1134
Db 1081 GPCSRERKKLCCQISDYCLKFNFPYSLEYVNCIKSBIKSPYKCFKSEGOSSM 1134

RESULT 6

Q8MT18_PLAFA

ID Q8MT18_PLAFA PRELIMINARY; PRT; 1137 AA.

AC Q8MT18;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Erythrocyte-binding protein.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Mayer D.G., Mu J.-B., Peng X., Su X.-Z., Miller L.H.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY099887; AAM51623.1; -; Genomic DNA.
DR GO: 0016021; C: integral to membrane; IEA.
DR GO: 0004872; F: receptor activity; IEA.
DR GO: 0009405; P: pathogenesis; IEA.
DR InterPro: IPR008602; Duff binding.
DR Pfam: PF05424; Duff binding; 1.
SQ SEQUENCE 1137 AA; 132293 MW; 43BDFAB02B0861D2 CRC64;

Query Match 93.8%; Score 6080; DB 2; Length 1137;
Best Local Similarity 99.6%; Pred. No. 3.8e-247;
Matches 1130; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKGVENLYFLIPLIFLYNVIRINESIIIGRTLYNRQDESSDISRVNSPELANNHKTNIYDS 60
Db 1 MKGVENLYFLIPLIFLYNVIRINESIIIGRTLYNRQDESSDISRVNSPELANNHKTNIYDS 60
Qy 61 DYEDVNNKLNINSFVENSVKKRSLSPINNKTSYDIIPPSYSYRNDKFNLSSENDNSG 120
Db 61 DYEDVNNKLNINSFVENSVKKRSLSPINNKTSYDIIPPSYSYRNDKFNLSSENDNSG 120
Qy 121 NTNSNFPANTSEISIGKNDKQYTFIQKRTFLFACGIRKRSIKWICRENSSEKITVCPDRK 180
Db 121 NTNSNFPANTSEISIGKNDKQYTFIQKRTFLFACGIRKRSIKWICRENSSEKITVCPDRK 180
Qy 181 IQLCVANFLNRLTMEKFKEIFLISVNTAEKLLYNKNEGKPSIFCNELRNSFSDFRSS 240
Db 181 IQLCVANFLNRLTMEKFKEIFLISVNTAEKLLYNKNEGKPSIFCNELRNSFSDFRSS 240
Qy 241 FIGDDMPGGTDRVKGINTKFSYDYKQKVEKLNKIKKEWKEKKNANLWNIHVMHKG 300
Db 241 FIGDDMPGGTDRVKGINTKFSYDYKQKVEKLNKIKKEWKEKKNANLWNIHVMHKG 300
Qy 301 NISKECAIIPAEFPQINLWIKENENFLMEKRLFLNIDKCVENKKEKYEACFGCRLPCS 360
Db 301 NISKECAIIPAEFPQINLWIKENENFLMEKRLFLNIDKCVENKKEKYEACFGCRLPCS 360
Qy 361 SYTSFMKSKTQMEVLNLYKKQNSGVDKNNFLNDFKNNKNDLDDFFKNEKEYDLDL 420
Db 361 SYTSFMKSKTQMEVLNLYKKQNSGVDKNNFLNDFKNNKNDLDDFFKNEKEYDLDL 420
Qy 421 CRYTATIIKSLFNGPAKNDVDIASQINVDNLRGFGCNYSKNEKSWNCTGTFNKPFGTC 480
Db 421 CRYTATIIKSLFNGPAKNDVDIASQINVDNLRGFGCNYSKNEKSWNCTGTFNKPFGTC 480
Qy 481 EPPRQTLCLGRYTLHRRGHEEDYKEHLLGASIIYEAQLLKYKKEKDNALCSIIQNSYA 540
Db 481 EPPRQTLCLGRYTLHRRGHEEDYKEHLLGASIIYEAQLLKYKKEKDNALCSIIQNSYA 540
Qy 541 DLADIIKGSDDIIKDYKQKMEENLKNVNDKKNRNEESLKIPEKRWMDENKENVKWSAV 600
Db 541 DLADIIKGSDDIIKDYKQKMEENLKNVNDKKNRNEESLKIPEKRWMDENKENVKWSAV 600
Qy 601 LKNEKTCQYDKFQKIPOFLWFKWGGDDFCERKEKIIYSPESFKVECKKCDENTCKN 660
Db 601 LKNEKTCQYDKFQKIPOFLWFKWGGDDFCERKEKIIYSPESFKVECKKCDENTCKN 660
Qy 661 KCSYKWKWIDLKSEYKQVDTYKDKNKKQYNDIDEVONKEANVYLKESKECKQVNF 720
Db 661 KCSYKWKWIDLKSEYKQVDTYKDKNKKQYNDIDEVONKEANVYLKESKECKQVNF 720
Qy 721 DKIFNESPNEVEDMCKKDEIKYLNEIKYPTKHDIIYDIDTFSDTFGDTGPISINANINE 780
Db 721 DKIFNESPNEVEDMCKKDEIKYLNEIKYPTKHDIIYDIDTFSDTFGDTGPISINANINE 780

```
QY 781 QOSGKDTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSSETRGILDINDPSVTNNV 840
Db 781 QOSGKDTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSSETRGILDINDPSVTNNV 840
QY 841 EVHDASTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSSETRGILDINDPSVTNNV 900
Db 841 EVHDASTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSSETRGILDINDPSVTNNV 900
QY 901 SLTIGQVPSDNTQNTYDSQNPHRDTPNALASLPSDDKINEIEGDFSRRSENGRGDTTS 960
Db 901 SLTIGQVPSDNTQNTYDSQNPHRDTPNALASLPSDDKINEIEGDFSRRSENGRGDTTS 960
QY 961 NTHDVRTNIVSERRVNSHDFIRNGMANNAHQYITQIENGLIIRQOESAGNSVNYKD 1020
Db 961 NTHDVRTNIVSERRVNSHDFIRNGMANNAHQYITQIENGLIIRQOESAGNSVNYKD 1020
QY 1021 NPKRSNFSSNDHKNIQYNSRDTKRVREIILSKQNKCNNEYSMEYCTYSDERNSSP 1080
Db 1021 NPKRSNFSSNDHKNIQYNSRDTKRVREIILSKQNKCNNEYSMEYCTYSDERNSSP 1080
QY 1081 GPCSRERKKLCCQISDYCLKYFNFSYIEYNYCIKSEIKSPYKCFKSEQSSII 1134
Db 1081 GPCSRERKKLCCQISDYCLKYFNFSYIEYNYCIKSEIKSPYKCFKSEQSSII 1134

RESULT 7
Q8MTI6_PLAFA
ID Q8MTI6_PLAFA PRELIMINARY; PRT; 1137 AA.
AC Q8MTI6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Erythrocyte-binding protein.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Mayer D.G., Mu J.-B., Feng X., Su X.-Z., Miller L.H.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY099889; AAM51625.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy binding.
DR Pfam; PF05424; Duffy binding; 1.
SQ SEQUENCE 1137 AA; 132348 MW; 30C28B9B376C76A7 CRC64;

Query Match 93.8%; Score 6077; DB 2; Length 1137;
Best Local Similarity 99.6%; Pred. No. 5.1e-247;
Matches 1129; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKGYFNLYFLPLFLVNLVIRINESIIIGRTLYNRQDESSDIIRVNSPELNNHKTNIYDS 60
Db 1 MKGYFNLYFLPLFLVNLVIRINESIIIGRTLYNRQDESSDIIRVNSPELNNHKTNIYDS 60

QY 61 DYEDVNNKLINSFVENKSVKKRSLSPFNKTKSYDIIPPSYSYRNDKFNSENEEDNSG 120
Db 61 DYEDVNNKLINSFVENKSVKKRSLSPFNKTKSYDIIPPSYSYRNDKFNSENEEDNSG 120

QY 121 NTSNNNFANTSEISIGKNTQYTFIQKTHLFCAGIKRKSIKWICRENSKITTVCVDRK 180
Db 121 NTSNNNFANTSEISIGKNTQYTFIQKTHLFCAGIKRKSIKWICRENSKITTVCVDRK 180

QY 181 IQLCVANFLASRLTMEKFEIPLISVNTAKLLYNKNEKDPISIFCNELRNSFSDPRSS 240
Db 181 IQLCVANFLASRLTMEKFEIPLISVNTAKLLYNKNEKDPISIFCNELRNSFSDPRSS 240

QY 241 FIGDDMDFGGNTDRVKGYINTKFSDDYYKKNVEKLNINKKEWKEKKNKANLWNHMI 300
Db 241 FIGDDMDFGGNTDRVKGYINTKFSDDYYKKNVEKLNINKKEWKEKKNKANLWNHMI 300

QY 301 NISKECAIIPAEBPQINLWIKWENNFLMEKRLFLNLIKDCVENKKEYACFGGCRLP 360
```

```
Db 301 NISKECAIIPAEBPQINLWIKWENNFLMEKRLFLNLIKDCVENKKEYACFGGCRLP 360
QY 361 SYTSFMKKSKTQMEVLTNLYKKKNSGVDKNNFLNDLFKKNKNDLDDFPKNEKEYDDLCD 420
Db 361 SYTSFMKKSKTQMEVLTNLYKKKNSGVDKNNFLNDLFKKNKNDLDDFPKNEKEYDDLCD 420
QY 421 CRYTATIISFLNGPAKNDVDIASQINVDLRLGFCGNYKSNNEKSNWCTGTFTNKPFGTC 480
Db 421 CRYTATIISFLNGPAKNDVDIASQINVDLRLGFCGNYKSNNEKSNWCTGTFTNKPFGTC 480
QY 481 EPPRRQTLCLGRTYLLHRGHEEDYKEHLIGASLYEAQLLYKYKEKDENALCSIIQNSYA 540
Db 481 EPPRRQTLCLGRTYLLHRGHEEDYKEHLIGASLYEAQLLYKYKEKDENALCSIIQNSYA 540
QY 541 DLADIISGSDIISKDYIGKKEENLNKVKNDKKNBESLKI FRKWKNDENKENVKWSAV 600
Db 541 DLADIISGSDIISKDYIGKKEENLNKVKNDKKNBESLKI FRKWKNDENKENVKWSAV 600
QY 601 LKNKETCKDYDFQKIPQFLRWFKEWGGDFCEKREKIKYSFESFKVECKKCCDENTCKN 660
Db 601 LKNKETCKDYDFQKIPQFLRWFKEWGGDFCEKREKIKYSFESFKVECKKCCDENTCKN 660
QY 661 KCSEYKKWIDLKSEYEKQVDKYTKDNKKNMYDNIDEVKNKEANVYLKEKSKCKOVNFD 720
Db 661 KCSEYKKWIDLKSEYEKQVDKYTKDNKKNMYDNIDEVKNKEANVYLKEKSKCKOVNFD 720
QY 721 DKIFNESPNEDMCKCDEIKYLNIEIKYPTKHDYIDITFSDTFGDGTPIISINANINE 780
Db 721 DKIFNESPNEDMCKCDEIKYLNIEIKYPTKHDYIDITFSDTFGDGTPIISINANINE 780
QY 781 QOSGKDTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSSETRGILDINDPSVTNNV 840
Db 781 QOSGKDTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSSETRGILDINDPSVTNNV 840
QY 841 EVHDASTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSSETRGILDINDPSVTNNV 900
Db 841 EVHDASTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSSETRGILDINDPSVTNNV 900
QY 901 SLTIGQVPSDNTQNTYDSQNPHRDTPNALASLPSDDKINEIEGDFSRRSENGRGDTTS 960
Db 901 SLTIGQVPSDNTQNTYDSQNPHRDTPNALASLPSDDKINEIEGDFSRRSENGRGDTTS 960
QY 961 NTHDVRTNIVSERRVNSHDFIRNGMANNAHQYITQIENGLIIRQOESAGNSVNYKD 1020
Db 961 NTHDVRTNIVSERRVNSHDFIRNGMANNAHQYITQIENGLIIRQOESAGNSVNYKD 1020
QY 1021 NPKRSNFSSNDHKNIQYNSRDTKRVREIILSKQNKCNNEYSMEYCTYSDERNSSP 1080
Db 1021 NPKRSNFSSNDHKNIQYNSRDTKRVREIILSKQNKCNNEYSMEYCTYSDERNSSP 1080
QY 1081 GPCSRERKKLCCQISDYCLKYFNFSYIEYNYCIKSEIKSPYKCFKSEQSSII 1134
Db 1081 GPCSRERKKLCCQISDYCLKYFNFSYIEYNYCIKSEIKSPYKCFKSEQSSII 1134

RESULT 8
Q8MTI7_PLAFA
ID Q8MTI7_PLAFA PRELIMINARY; PRT; 1137 AA.
AC Q8MTI7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Erythrocyte-binding protein.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Mayer D.G., Mu J.-B., Feng X., Su X.-Z., Miller L.H.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY099888; AAM51624.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
```

DR GO:0004872; P:receptor activity; IEA.
DR GO:0009405; P:pathogenesis; IEA.
DR InterPro: IPR008602; Duffy_binding.
DR Pfam: PF05424; Duffy_binding; 1.
SQ SEQUENCE 1137 AA; 132320 MW; FFC3C136F697A88A CRC64;

Query Match 93.8%; Score 6077; DB 2; Length 1137;
Best Local Similarity 99.6%; Pred. No. 5.1e-247;
Matches 1129; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKGFNYFLIPLIFLYNVRINESIIIGRTLYNRQDESSDIISRVNSPELNNHKTNIYDS 60
Db 1 MKGFNYFLIPLIFLYNVRINESIIIGRTLYNRQDESSDIISRVNSPELNNHKTNIYDS 60

Qy 61 DYEDVNNKLINSFVENKSVKKRSLSPINNKTSDYDIIPPSYRNDRKNSLSENEONSG 120
Db 61 DYEDVNNKLINSFVENKSVKKRSLSPINNKTSDYDIIPPSYRNDRKNSLSENEONSG 120

Qy 121 NTNSNNFANTSEISIGKDNKQYTFIQKRTLFAAGIKRKSIIKWCIRENSEKITVCVPRK 180
Db 121 NTNSNNFANTSEISIGKDNKQYTFIQKRTLFAAGIKRKSIIKWCIRENSEKITVCVPRK 180

Qy 181 IQLCVANPLNSRLTMEKFKEIFLISVNTAKLLYNKNEGKDPISFCNELRNSPSPRSS 240
Db 181 IQLCVANPLNSRLTMEKFKEIFLISVNTAKLLYNKNEGKDPISFCNELRNSPSPRSS 240

Qy 241 FIGDDMPGGNTDRVKGIVNTKPSDYKKNVEKLNNIKKEMWKNKANLWNHMIVNHKG 300
Db 241 FIGDDMPGGNTDRVKGIVNTKPSDYKKNVEKLNNIKKEMWKNKANLWNHMIVNHKG 300

Qy 301 NISKECAIIPAEPOINLWIKEMWENFLMEKRLFLNIDKCKVENKYYEACFGGCRLPSCS 360
Db 301 NISKECAIIPAEPOINLWIKEMWENFLMEKRLFLNIDKCKVENKYYEACFGGCRLPSCS 360

Qy 361 SYTSFMKSKTQMEVNLNLYKKNSGVDKNNFLNDFKCKNNKNDLDDFFKNEKEYDDLCD 420
Db 361 SYTSFMKSKTQMEVNLNLYKKNSGVDKNNFLNDFKCKNNKNDLDDFFKNEKEYDDLCD 420

Qy 421 CRYTATIIKSPLANGPAKNDVDIASQINVDLRGFCNYSKNSKSWNCTGFTTNKPGTC 480
Db 421 CRYTATIIKSPLANGPAKNDVDIASQINVDLRGFCNYSKNSKSWNCTGFTTNKPGTC 480

Qy 481 EPPRRQTLCLGRTYLLHGRHEEDYKHLGASIIYEAQLLYKKEKDNALCSIIQNSYA 540
Db 481 EPPRRQTLCLGRTYLLHGRHEEDYKHLGASIIYEAQLLYKKEKDNALCSIIQNSYA 540

Qy 541 DLADIIKSGDIIIDYKGGKBEENLVNKKDKRNEESLKIIPREKWDENKENVWVMSAV 600
Db 541 DLADIIKSGDIIIDYKGGKBEENLVNKKDKRNEESLKIIPREKWDENKENVWVMSAV 600

Qy 601 LKNKETCKDYKPKQIKIQLFWFKWGGDDFCERKEKIKIYSPESPKVECKKCDENTCKN 660
Db 601 LKNKETCKDYKPKQIKIQLFWFKWGGDDFCERKEKIKIYSPESPKVECKKCDENTCKN 660

Qy 661 KCSYKWKWIDLKSEYKQVDKYTKDKNNKMYDNIDEVNKEANVYLKESKECKDVNFD 720
Db 661 KCSYKWKWIDLKSEYKQVDKYTKDKNNKMYDNIDEVNKEANVYLKESKECKDVNFD 720

Qy 721 DKIFNESPNEYDMCKKDEIKYLNEIKYPTKTHDIYDIDTFSDTFDGTGPISINANINE 780
Db 721 DKIFNESPNEYDMCKKDEIKYLNEIKYPTKTHDIYDIDTFSDTFDGTGPISINANINE 780

Qy 781 QOSGKDTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSETRGILDINDPSVTNNVN 840
Db 781 QOSGKDTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSETRGILDINDPSVTNNVN 840

Qy 841 EVHDASNTQGSVNTSDITNGHSSSLNRTTNAQDIKIGRSGNQSDNOENSHSSDMSG 900
Db 841 EVHDASNTQGSVNTSDITNGHSSSLNRTTNAQDIKIGRSGNQSDNOENSHSSDMSG 900

Qy 901 SLTTIGQVPSDNTQNTYDSQNPHRDTPNALASLPDDKINEIEGFDSSRSSENGRDTTS 960
Db 901 SLTTIGQVPSDNTQNTYDSQNPHRDTPNALASLPDDKINEIEGFDSSRSSENGRDTTS 960

Qy 961 NTHDVRRTNIVSRRRVNSHDFIRNGMANNNAHQQYITQIENNGIIRGOESAGNSVNYKD 1020
Db 961 NTHDVRRTNIVSRRRVNSHDFIRNGMANNNAHQQYITQIENNGIIRGOESAGNSVNYKD 1020

Qy 1021 NPKRSNFSSENHKKONQYNSRDTKRVREIILKSKONKCNNEYSMEYCTYSDERNSSP 1080
Db 1021 NPKRSNFSSENHKKONQYNSRDTKRVREIILKSKONKCNNEYSMEYCTYSDERNSSP 1080

Qy 1081 GPCSRERKKLCCQISDYCLKYFNFYISYIYNYNCIKSEIKSPYKCFKSEGOSSI 1134
Db 1081 GPCSRERKKLCCQISDYCLKYFNFYISYIYNYNCIKSEIKSPYKCFKSEGOSSM 1134

RESULT 9
QSEK8_PLARE PRELIMINARY; PRT; 1282 AA.
AC QSEK8;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Erythrocyte invasion ligand BABBL/EEA-140 (Fragment).
OS Plasmodium reichenowi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5854;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Rayner J.C., Huber C.S., Barnwell J.W.;
RT "Conservation and divergence in erythrocyte invasion ligands:
Mol. Biochem. Parasitol. 138:243-247 (2004).
DR EMBL: AV572433; AAT7188.1; -; Genomic_DNA.
DR InterPro: IPR008602; Duffy_binding.
DR Pfam: PF05424; Duffy_binding; 1.
FT NON TER 1
SQ SEQUENCE 1282 AA; 148469 MW; 53D601B4AE7CEBCC CRC64;

Query Match 86.5%; Score 5606.5; DB 2; Length 1282;
Best Local Similarity 82.0%; Pred. No. 3.2e-227;
Matches 1058; Conservative 50; Mismatches 90; Indels 93; Gaps 2;

Qy 4 YFNIYFLIPLIFLYNVRINESIIIGRTLYNRQDESSDIISRVNSPELNNHKTNIYDSYE 63
Db 1 YFNIYFLIPLIFLYNVRINESIIIGRTLYNRQDESSDIISRVNSPELNNHKTNIYDSYE 60

Qy 64 DVNNKLINSFVENKSVKKRSLSPINNKTSDYDIIPPSYRNDRKNSLSENEONSGNTN 123
Db 61 DVNNKLINSFVENKSVKKRSLSPINNKTSDYDIIPPSYRNDRKNSLSENEONSGNTN 120

Qy 124 SNFPANTSEISIGKDNKQYTFIQKRTLFAAGIKRKSIIKWCIRENSEKITVCVPRKIQ 183
Db 121 SNFPANTSEISIGKDNKQYTFIQKRTLFAAGIKRKSIIKWCIRENSEKITVCVPRKIQ 180

Qy 184 CVANFLNSRLTMEKFKEIFLISVNTAKLLYNKNEGKDPISFCNELRNSPSPRSSFIG 243
Db 181 CVANFLNSRLTMEKFKEIFLISVNTAKLLYNKNEGKDPISFCNELRNSPSPRSSFIG 240

Qy 244 DDMDFGNTDRVKGIVNTKPSDYKKNVEKLNNIKKEMWKNKANLWNHMIVNHKNIS 303
Db 241 DDMDFGNTDRVKGIVNTKPSDYKKNVEKLNNIKKEMWKNKANLWNHMIVNHKNIS 300

Qy 304 KECAIIPAEPOINLWIKEMWENFLMEKRLFLNIDKCKVENKYYEACFGGCRLPSCSYT 363
Db 301 KECAIIPAEPOINLWIKEMWENFLMEKRLFLNIDKCKVENKYYEACFGGCRLPSCSYT 360

Qy 364 SFMKSKTQMEVNLNLYKKNSGVDKNNFLNDFKCKNNKNDLDDFFKNEKEYDDLCDRY 423
Db 361 SFMKSKTQMEVNLNLYKKNSGVDKNNFLNDFKCKNNKNDLDDFFKNEKEYDDLCDRY 420

Qy 424 TATIISKFLNGPAKNDVDIASQINVDLRGFCNYSKNSKSWNCTGFTTNKPGTCVPP 483
Db 421 TATIISKFLNGPAKNDVDIASQINVDLRGFCNYSKNSKSWNCTGFTTNKPGTCVPP 480

DE (Fragment).
OS Name=eba-140;
GN Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]

NUCLEOTIDE SEQUENCE.

RA Mayer D.G., Mu J.-B., Feng X., Su X.-Z., Miller L.H.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

[2]

RP MEDLINE=22588500; PubMed=12702678;
RX Baum J., Thomas A.W., Conway D.J.;

RA "Evidence for diversifying selection on erythrocyte-binding antigens
of Plasmodium falciparum and P. vivax.";
RT Genetics 163:1327-1336(2003).

DR EMBL; AF507981; AAM45246.1; -; Genomic DNA.
DR EMBL; AF507983; AAM45248.1; -; Genomic DNA.
DR EMBL; AF507984; AAM45249.1; -; Genomic DNA.
DR EMBL; AF507985; AAM45250.1; -; Genomic DNA.
DR EMBL; AJ438833; CAD27558.1; -; Genomic DNA.
DR EMBL; AJ438834; CAD27559.1; -; Genomic DNA.
DR EMBL; AJ438839; CAD27564.1; -; Genomic DNA.
DR EMBL; AJ438849; CAD27574.1; -; Genomic DNA.
DR EMBL; AF507982; AAM45247.1; -; Genomic DNA.

FT NON TER 1 616
FT NON TER 616 616

SQ SEQUENCE 616 AA; 73456 MW; 6FDE19BPEC982587 CRC64;

Query Match 52.0%; Score 3368; DB 2; Length 616;

Best Local Similarity 99.8%; Pred. No. 1.2e-133;

Matches 615; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 141 QYTFIQKRTFLFACGIGKRSIKWICRENSKIVCVDPDKIQLCVANFLNSRLTMEKFK 200
Db 1 QYTFIQKRTFLFACGIGKRSIKWICRENSKIVCVDPDKIQLCVANFLNSRLTMEKFK 60

Qy 201 EIFLISVNTAKLLYNKNEGKPSIFCNELNSFSDFRSFPGDMDFGGNTDRVKGIN 260
Db 61 EIFLISVNTAKLLYNKNEGKPSIFCNELNSFSDFRSFPGDMDFGGNTDRVKGIN 120

Qy 261 TKFSDDYKKNVEKLNKIKKEWKEKNKANLWNHMIWNHKNISKECAIIPAEFPQINLWI 320
Db 121 KKFSDYYKKNVEKLNKIKKEWKEKNKANLWNHMIWNHKNISKECAIIPAEFPQINLWI 180

Qy 321 KEWNEFLMEKRLFLNIKDKCVENKYYEACFGCRLPCSSYTSFMKSKTQMEVLNLY 380
Db 181 KEWNEFLMEKRLFLNIKDKCVENKYYEACFGCRLPCSSYTSFMKSKTQMEVLNLY 240

Qy 381 KKNNSGVDKNNFLNDLPKNNKNDLDDFPKNEKEYDLDLCRYTATIIKSGFLNGPAKNDV 440
Db 241 KKNNSGVDKNNFLNDLPKNNKNDLDDFPKNEKEYDLDLCRYTATIIKSGFLNGPAKNDV 300

Qy 441 DIASQINVDLARGFCNYSKNSNKNWCTGTFTNKPFGTCPPRROTLCGLRTYLLHRGH 500
Db 301 DIASQINVDLARGFCNYSKNSNKNWCTGTFTNKPFGTCPPRROTLCGLRTYLLHRGH 360

Qy 501 EEDYKEHLLGASIEAQLLYKYKKEKDNALCSIIQNSYADLADIIGSDIIKDYGGKM 560
Db 361 EEDYKEHLLGASIEAQLLYKYKKEKDNALCSIIQNSYADLADIIGSDIIKDYGGKM 420

Qy 561 EENLNKYNKDKRNEESIKIFREKWDENKENVKMSAVLNKKTCKDYDKFKQIKPOFL 620
Db 421 EENLNKYNKDKRNEESIKIFREKWDENKENVKMSAVLNKKTCKDYDKFKQIKPOFL 480

Qy 621 RWFKEWDDDFCEKEKEKIEYFESFKVECKKCDENTCKNCSYKKKWLIDAKSEYKQV 680
Db 481 RWFKEWDDDFCEKEKEKIEYFESFKVECKKCDENTCKNCSYKKKWLIDAKSEYKQV 540

Qy 681 DKYTKDKNNKMYDNIDEVKNKENVYLKESKECKVNFDDKIPNESPNYEDMCKKCD 740
Db 541 DKYTKDKNNKMYDNIDEVKNKENVYLKESKECKVNFDDKIPNESPNYEDMCKKCD 600

Qy 741 IKYLNEIKYPTKXHI 756
Db 601 IKYLNEIKYPTKXHI 616

RESULT 12

Q8T9L8 PLAF

ID Q8T9L8 PLAF PRELIMINARY; PRT; 616 AA.

AC Q8T9L8;

DT 01-JUN-2002 (T-EMBLrel. 21, Created)

DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)

DT 01-FEB-2005 (T-EMBLrel. 29, Last annotation update)

DE Erythrocyte binding protein 2 (BAEBL protein) (Fragment).

GN Name=baebl;

OS Plasmodium falciparum.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=5833;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=Oak Knoll;

RX MEDLINE=21674692; PubMed=11814568; DOI=10.1016/S0166-6851(01)00428-5;

RA Narum D.L., Fuhrmann S.R., Luu T., Sim B.K.;

RT "A novel Plasmodium falciparum erythrocyte binding protein-2

(BBP2/BAEBL) involved in erythrocyte receptor binding.";

RL Mol. Biochem. Parasitol. 119:159-168(2002).

[2]

NUCLEOTIDE SEQUENCE

RA Mayer D.G., Mu J.-B., Feng X., Su X.-Z., Miller L.H.;

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY052371; AAL23761.1; -; Genomic DNA.

DR EMBL; AF507979; AAM45244.1; -; Genomic DNA.

DR EMBL; AF507980; AAM45245.1; -; Genomic DNA.

DR EMBL; AF507978; AAM45243.1; -; Genomic DNA.

FT NON TER 1 616

FT NON TER 616 616

SQ SEQUENCE 616 AA; 73470 MW; 419A89BAB4AA1F8F CRC64;

Query Match 52.0%; Score 3367; DB 2; Length 616;

Best Local Similarity 99.7%; Pred. No. 1.4e-133;

Matches 614; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 141 QYTFIQKRTFLFACGIGKRSIKWICRENSKIVCVDPDKIQLCVANFLNSRLTMEKFK 200
Db 1 QYTFIQKRTFLFACGIGKRSIKWICRENSKIVCVDPDKIQLCVANFLNSRLTMEKFK 60

Qy 201 EIFLISVNTAKLLYNKNEGKPSIFCNELNSFSDFRSFPGDMDFGGNTDRVKGIN 260
Db 61 EIFLISVNTAKLLYNKNEGKPSIFCNELNSFSDFRSFPGDMDFGGNTDRVKGIN 120

Qy 261 TKFSDDYKKNVEKLNKIKKEWKEKNKANLWNHMIWNHKNISKECAIIPAEFPQINLWI 320
Db 121 KKFSDYYKKNVEKLNKIKKEWKEKNKANLWNHMIWNHKNISKECAIIPAEFPQINLWI 180

Qy 321 KEWNEFLMEKRLFLNIKDKCVENKYYEACFGCRLPCSSYTSFMKSKTQMEVLNLY 380
Db 181 KEWNEFLMEKRLFLNIKDKCVENKYYEACFGCRLPCSSYTSFMKSKTQMEVLNLY 240

Qy 381 KKNNSGVDKNNFLNDLPKNNKNDLDDFPKNEKEYDLDLCRYTATIIKSGFLNGPAKNDV 440
Db 241 KKNNSGVDKNNFLNDLPKNNKNDLDDFPKNEKEYDLDLCRYTATIIKSGFLNGPAKNDV 300

Qy 441 DIASQINVDLARGFCNYSKNSNKNWCTGTFTNKPFGTCPPRROTLCGLRTYLLHRGH 500
Db 301 DIASQINVDLARGFCNYSKNSNKNWCTGTFTNKPFGTCPPRROTLCGLRTYLLHRGH 360

Qy 501 EEDYKEHLLGASIEAQLLYKYKKEKDNALCSIIQNSYADLADIIGSDIIKDYGGKM 560
Db 361 EEDYKEHLLGASIEAQLLYKYKKEKDNALCSIIQNSYADLADIIGSDIIKDYGGKM 420

Qy 561 EENLNKYNKDKRNEESIKIFREKWDENKENVKMSAVLNKKTCKDYDKFKQIKPOFL 620
Db 421 EENLNKYNKDKRNEESIKIFREKWDENKENVKMSAVLNKKTCKDYDKFKQIKPOFL 480

```
QY 621 RWFKEGDDDFCEKREKIYFSEFVKVECKKDCDENTCKNKCSEYKWKWIDLKSEYEKQV 680
Db 481 RWFKEGDDDFCEKREKIYFSEFVKVECKKDCDENTCKNKCSEYKWKWIDLKSEYEKQV 540
QY 681 DKYTKDKNKKMYDNIDEVKNKEANVYLKEKCKDVNFDKIFNESPNEDYEDMCKKCD 740
Db 541 DKYTKDKNKKMYDNIDEVKNKEANVYLKEKCKDVNFDKIFNESPNEDYEDMCKKCD 600
QY 741 IKYLNKIKYPTKXDI 756
Db 601 IKYLNKIKYPTKXDI 616

RESULT 13
Q81049_PLAFA PRELIMINARY; PRT; 616 AA.
AC Q81049;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Erythrocyte binding antigen region II (Fragment).
GN Name=eba-140;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22588500; PubMed=12702678;
RA Baum J., Thomas A.W., Conway D.J.;
RT "Evidence for diversifying selection on erythrocyte-binding antigens
of Plasmodium falciparum and P. vivax.";
RL Genetics 163:1327-1336(2003).
DR EMBL; AJ438832; CAD27557.1; -; Genomic_DNA.
DR EMBL; AJ438836; CAD27561.1; -; Genomic_DNA.
DR EMBL; AJ438850; CAD27575.1; -; Genomic_DNA.
FT NON_TER 1
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 616 AA; 73483 MW; F6DE19BFEC9FC344 CRC64;

Query Match 51.9%; Score 3365; DB 2; Length 616;
Best Local Similarity 99.7%; Pred. No. 1.7e-133;
Matches 614; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 141 QYTFIQKRTLFCAGIKRKSIKWI CRENSEKITVCVPDRKIQLCVANFLNSRLTMEKPK 200
Db 1 QYTFIQKRTLFCAGIKRKSIKWI CRENSEKITVCVPDRKIQLCVANFLNSRLTMEKPK 60

QY 201 EIFLISVNTAEKLLYNKNEGKPSIFCNELNSFSDFRSFSGDMDPGGNTDRVKGYIN 260
Db 61 EIFLISVNTAEKLLYNKNEGKPSIFCNELNSFSDFRSFSGDMDPGGNTDRVKGYIN 120

QY 261 TKFSDDYKKNVEKLNLIKKWEKKNKANLWNHMI VNHKGNISKECAII PAEPPQINLWI 320
Db 121 KTFSDYYKKNVEKLNLIKKWEKKNKANLWNHMI VNHKGNISKECAII PAEPPQINLWI 180

QY 321 KENNENFLMEKKRLFLNI KDCVENKKEYEACFGGCRLLPCSSYTSFMKSKTQMEVLNLY 380
Db 181 KENNENFLMEKKRLFLNI KDCVENKKEYEACFGGCRLLPCSSYTSFMKSKTQMEVLNLY 240

QY 381 KKNKSGVDKNNFLNDFKKNKNKNDLDDFFKNEKEYDDLCDCRYTATIIKSFNGPAKNDV 440
Db 241 KKNKSGVDKNNFLNDFKKNKNKNDLDDFFKNEKEYDDLCDCRYTATIIKSFNGPAKNDV 300

QY 441 DIASQINVDLRGFCNYSKSNKSNWCTGTFTNKFPGTCEPPRRQTLCLGRYVLLHRGH 500
Db 301 DIASQINVDLRGFCNYSKSNKSNWCTGTFTNKFPGTCEPPRRQTLCLGRYVLLHRGH 360

QY 501 EEDYKEHLLGASIEYAQLLKYKKEKDNALCSIIQNSYADLADI IKGSDIIKDYGGK 560
Db 361 EEDYKEHLLGASIEYAQLLKYKKEKDNALCSIIQNSYADLADI IKGSDIIKDYGGK 420

QY 561 EENLKNVKDKKRNESLSIKFREKWDENKENVKMSAVLNKNETCKDYDKFKQIPQFL 620
Db 421 EENLKNVKDKKRNESLSIKFREKWDENKENVKMSAVLNKNETCKDYDKFKQIPQFL 480
QY 621 RWFKEGDDDFCEKREKIYFSEFVKVECKKDCDENTCKNKCSEYKWKWIDLKSEYEKQV 680
```

```
Db 421 EENLKNVKDKKRNESLSIKFREKWDENKENVKMSAVLNKNETCKDYDKFKQIPQFL 480
QY 621 RWFKEGDDDFCEKREKIYFSEFVKVECKKDCDENTCKNKCSEYKWKWIDLKSEYEKQV 680
Db 481 RWFKEGDDDFCEKREKIYFSEFVKVECKKDCDENTCKNKCSEYKWKWIDLKSEYEKQV 540
QY 681 DKYTKDKNKKMYDNIDEVKNKEANVYLKEKCKDVNFDKIFNESPNEDYEDMCKKCD 740
Db 541 DKYTKDKNKKMYDNIDEVKNKEANVYLKEKCKDVNFDKIFNESPNEDYEDMCKKCD 600
QY 741 IKYLNKIKYPTKXDI 756
Db 601 IKYLNKIKYPTKXDI 616

RESULT 14
Q8MM74_PLAFA PRELIMINARY; PRT; 616 AA.
AC Q8MM74;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE BAEBL protein (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Mayer D.G., Mu J.-B., Feng X., Su X.-Z., Miller L.H.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF507989; AAM45254.1; -; Genomic_DNA.
DR EMBL; AF507990; AAM45255.1; -; Genomic_DNA.
FT NON_TER 1
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 616 AA; 73525 MW; 2C526A348C4E7771 CRC64;

Query Match 51.9%; Score 3364; DB 2; Length 616;
Best Local Similarity 99.5%; Pred. No. 1.8e-133;
Matches 613; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 141 QYTFIQKRTLFCAGIKRKSIKWI CRENSEKITVCVPDRKIQLCVANFLNSRLTMEKPK 200
Db 1 QYTFIQKRTLFCAGIKRKSIKWI CRENSEKITVCVPDRKIQLCVANFLNSRLTMEKPK 60

QY 201 EIFLISVNTAEKLLYNKNEGKPSIFCNELNSFSDFRSFSGDMDPGGNTDRVKGYIN 260
Db 61 EIFLISVNTAEKLLYNKNEGKPSIFCNELNSFSDFRSFSGDMDPGGNTDRVKGYIN 120

QY 261 TKFSDDYKKNVEKLNLIKKWEKKNKANLWNHMI VNHKGNISKECAII PAEPPQINLWI 320
Db 121 RKFSDYYKKNVEKLNLIKKWEKKNKANLWNHMI VNHKGNISKECAII PAEPPQINLWI 180

QY 321 KENNENFLMEKKRLFLNI KDCVENKKEYEACFGGCRLLPCSSYTSFMKSKTQMEVLNLY 380
Db 181 KENNENFLMEKKRLFLNI KDCVENKKEYEACFGGCRLLPCSSYTSFMKSKTQMEVLNLY 240

QY 381 KKNKSGVDKNNFLNDFKKNKNKNDLDDFFKNEKEYDDLCDCRYTATIIKSFNGPAKNDV 440
Db 241 KKNKSGVDKNNFLNDFKKNKNKNDLDDFFKNEKEYDDLCDCRYTATIIKSFNGPAKNDV 300

QY 441 DIASQINVDLRGFCNYSKSNKSNWCTGTFTNKFPGTCEPPRRQTLCLGRYVLLHRGH 500
Db 301 DIASQINVDLRGFCNYSKSNKSNWCTGTFTNKFPGTCEPPRRQTLCLGRYVLLHRGH 360

QY 501 EEDYKEHLLGASIEYAQLLKYKKEKDNALCSIIQNSYADLADI IKGSDIIKDYGGK 560
Db 361 EEDYKEHLLGASIEYAQLLKYKKEKDNALCSIIQNSYADLADI IKGSDIIKDYGGK 420

QY 561 EENLKNVKDKKRNESLSIKFREKWDENKENVKMSAVLNKNETCKDYDKFKQIPQFL 620
Db 421 EENLKNVKDKKRNESLSIKFREKWDENKENVKMSAVLNKNETCKDYDKFKQIPQFL 480
QY 621 RWFKEGDDDFCEKREKIYFSEFVKVECKKDCDENTCKNKCSEYKWKWIDLKSEYEKQV 680
```

Db 481 RWFKEWGDFFCEKKEKIYSPESFKVECKKDCDENTCKNCKSEYKWKWIDLKSEYKQV 540
 Qy 681 DKYTKDKNKKMYDNIDVKNKEANVYLKESKECKDVNFDKIKFNESPNEDMCKKDE 740
 Db 541 DKYTKDKNKKMYDNIDVKNKEANVYLKESKECKDVNFDKIKFNESPNEDMCKKDE 600
 Qy 741 IKYLNEIKYPTKXKDI 756
 Db 601 IKYLNEIKYPTKXKDI 616

RESULT 15

ID Q8T9N3 PLAFPA PRELIMINARY; PRT; 616 AA.
 AC Q8T9N3; OSMM01;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Erythrocyte binding protein-2 (BABBL protein) (Erythrocyte binding
 antigen region II) (Fragment).
 GN Name=BBP2; Synonyms=eba-140;
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=3D7;
 RA MEDLINE=21674692; PubMed=11814568; DOI=10.1016/S0166-6851(01)00428-5;
 RA Narum D.L., Fuhrmann S.R., Luu T., Sim B.K.;
 RT "A novel Plasmodium falciparum erythrocyte binding protein-2
 (BBP2/BABBL) involved in erythrocyte receptor binding.";
 RL Mol. Biochem. Parasitol. 119:159-168(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Mayer D.G., Mu J.-B., Feng X., Su X.-Z., Miller L.H.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]

NUCLEOTIDE SEQUENCE

RP MEDLINE=22588500; PubMed=12702678;
 RA Baum J., Thomas A.W., Conway D.J.;
 RT "Evidence for diversifying selection on erythrocyte-binding antigens
 of Plasmodium falciparum and P. vivax.";
 RL Genetics 163:1327-1336(2003).
 DR EMBL; AY030048; AAK55484.1; -; Genomic DNA.
 DR EMBL; AF507988; AAM45253.1; -; Genomic DNA.
 DR EMBL; AJ438830; CAD27555.1; -; Genomic DNA.
 DR EMBL; AJ438831; CAD27556.1; -; Genomic DNA.
 DR EMBL; AJ438837; CAD27562.1; -; Genomic DNA.
 DR EMBL; AJ438838; CAD27563.1; -; Genomic DNA.
 DR EMBL; AJ438842; CAD27567.1; -; Genomic DNA.
 DR EMBL; AJ438844; CAD27569.1; -; Genomic DNA.
 DR EMBL; AJ438845; CAD27570.1; -; Genomic DNA.
 DR EMBL; AJ438846; CAD27571.1; -; Genomic DNA.
 DR EMBL; AJ438848; CAD27573.1; -; Genomic DNA.
 DR EMBL; AJ438851; CAD27576.1; -; Genomic DNA.
 DR EMBL; AJ438852; CAD27577.1; -; Genomic DNA.
 DR EMBL; AJ438853; CAD27578.1; -; Genomic DNA.
 DR EMBL; AF507986; AAM45251.1; -; Genomic DNA.
 DR EMBL; AF507987; AAM45252.1; -; Genomic DNA.
 FT NON TER 1
 FT NON TER 1
 SQ SEQUENCE 616 AA; 73497 MW; DB9A9BAB4ADF94C CRC64;

Query Match 51.9%; Score 3364; DB 2; Length 616;
 Best Local Similarity 99.5%; Pred. No. 1.8e-133;
 Matches 613; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 141 QYTFIQKRTLFAAGIKRKSIKWKICRENSKIKITCVDPDRKIQLCVANFLNRLTMEKFK 200
 Db 1 QYTFIQKRTLFAAGIKRKSIKWKICRENSKIKITCVDPDRKIQLCVANFLNRLTMEKFK 60
 Qy 201 EIFLISVNTAEKLLYNKNEGKDPISFCNELRNSPDSFRSSFIGDMDFGGNTDRVKGYN 260

Db 61 EIFLISVNTAEKLLYNKNEGKDPISFCNELRNSPDSFRSSFIGDMDFGGNTDRVKGYN 120
 Qy 261 TKPSDYKKEKNVEKLNANI KKEWMEKNKANLWNHMI VNHKGNISKECAIIPAEBPQINLWI 320
 Db 121 KKFSDYYKKEKNVEKLNANI KKEWMEKNKANLWNHMI VNHKGNISKECAIIPAEBPQINLWI 180
 Qy 321 KENNENFLMEKRLFLNI KDKCVENKKYEACFGGCRLPCCSSYTSFMKSKTKQMEVLTNLY 380
 Db 181 KENNENFLMEKRLFLNI KDKCVENKKYEACFGGCRLPCCSSYTSFMKSKTKQMEVLTNLY 240
 Qy 381 KKNKSGVDKNNFLNDL FKNNKNDLDDF FKNKEKYDDL CDCRYTATII KSFINGPAKNDV 440
 Db 241 KKNKSGVDKNNFLNDL FKNNKNDLDDF FKNKEKYDDL CDCRYTATII KSFINGPAKNDV 300
 Qy 441 DIASQINVDLRGFCGNYKSNNEKSWNCTGFTNKFPGTCBPPTROTLCIGRTYLLHRGH 500
 Db 301 DIASQINVDLRGFCGNYKSNNEKSWNCTGFTNKFPGTCBPPTROTLCIGRTYLLHRGH 360
 Qy 501 BEDYKEHLGASIVYEAOLLYKYKEDENALCSIIQNSYADLADI IKGSDI IKDYYGKQM 560
 Db 361 BEDYKEHLGASIVYEAOLLYKYKEDENALCSIIQNSYADLADI IKGSDI IKDYYGKQM 420
 Qy 561 EENLNKYNKDKGRNEESLKI FKEKWMDENKENVKMSAVLNKNETCKDYDKFKI PQFL 620
 Db 421 EENLNKYNKDKGRNEESLKI FKEKWMDENKENVKMSAVLNKNETCKDYDKFKI PQFL 480
 Qy 621 RWPKEWGDFFCEKKEKIYSPESFKVECKKDCDENTCKNCKSEYKWKWIDLKSEYKQV 680
 Db 481 RWPKEWGDFFCEKKEKIYSPESFKVECKKDCDENTCKNCKSEYKWKWIDLKSEYKQV 540
 Qy 681 DKYTKDKNKKMYDNIDVKNKEANVYLKESKECKDVNFDKIKFNESPNEDMCKKDE 740
 Db 541 DKYTKDKNKKMYDNIDVKNKEANVYLKESKECKDVNFDKIKFNESPNEDMCKKDE 600
 Qy 741 IKYLNEIKYPTKXKDI 756
 Db 601 IKYLNEIKYPTKXKDI 616

RESULT 16

ID Q8I787 PLAFPA PRELIMINARY; PRT; 616 AA.
 AC Q8I787;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Erythrocyte binding antigen region II (Fragment).
 GN Name=eba-140;
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA MEDLINE=22588500; PubMed=12702678;
 RA Baum J., Thomas A.W., Conway D.J.;
 RT "Evidence for diversifying selection on erythrocyte-binding antigens
 of Plasmodium falciparum and P. vivax.";
 RL Genetics 163:1327-1336(2003).
 DR EMBL; AJ438840; CAD27565.1; -; Genomic DNA.
 FT NON TER 1
 FT NON TER 1
 SQ SEQUENCE 616 AA; 73442 MW; C50CE1C56E605F54 CRC64;

Query Match 51.9%; Score 3363; DB 2; Length 616;
 Best Local Similarity 99.7%; Pred. No. 2e-133;
 Matches 614; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 141 QYTFIQKRTLFAAGIKRKSIKWKICRENSKIKITCVDPDRKIQLCVANFLNRLTMEKFK 200
 Db 1 QYTFIQKRTLFAAGIKRKSIKWKICRENSKIKITCVDPDRKIQLCVANFLNRLTMEKFK 60
 Qy 201 EIFLISVNTAEKLLYNKNEGKDPISFCNELRNSPDSFRSSFIGDMDFGGNTDRVKGYN 260


```

Db 61 EIFLISVNTAEKLLYNKNEGKPSIFCNELRNSFSDFRSGFIQDDMDFGNTDRVKGYN 120
Qy 261 TKFSDDYKKEKVKELNNIKKEWMEKNKANLWNHMIWNHKGNIISKECAIIPAEBPQINLWI 320
Db 121 KFSDDYKKEKVKELNNIKKEWMEKNKANLWNHMIWNHKGNIISKECAIIPAEBPQINLWI 180
Qy 321 KEWNEFLMEKRLFLNIKDCKVENKYEACFGGCRLPCCSSYTSFMKKSQTQMEVLTNLY 380
Db 181 KEWNEFLMEKRLFLNIKDCKVENKYEACFGGCRLPCCSSYTSFMKKSQTQMEVLTNLY 240
Qy 381 KKNKSGVDKNNFLNDLFKKNKNKNDLDDFFKNEKEYDDLCDCRYTATTIISKFLNGPAKNDV 440
Db 241 KKNKSGVDKNNFLNDLFKKNKNKNDLDDFFKNEKEYDDLCDCRYTATTIISKFLNGPAKNDV 300
Qy 441 DIASQINVNDLRGFGCNYKSNNEKSNWCTGTFTNKPFGTCEPPTROTLCIGRTYLLHRGH 500
Db 301 DIASQINVNDLRGFGCNYKSNNEKSNWCTGTFTNKPFGTCEPPTROTLCIGRTYLLHRGH 360
Qy 501 EEDYKEHLLGASIEAQLLYKYKEKDENALCSIIQNSYADLADIIGSDIISKDYGGKQM 560
Db 361 EEDYKEHLLGASIEAQLLYKYKEKDENALCSIIQNSYADLADIIGSDIISKDYGGKQM 420
Qy 561 EENLNKVNKDKRNEESLKIIFREKWDENKENVKMSAVLNKNETCKDYDKFKQIPQFL 620
Db 421 EENLNKVNKDKRNEESLKIIFREKWDENKENVKMSAVLNKNETCKDYDKFKQIPQFL 480
Qy 621 RWFKEWGGDFCEKREKIYSFESFKVECKKDCDENTCKNCKSEYKMWIDLKSEYEKQV 680
Db 481 RWFKEWGGDFCEKREKIYSFESFKVECKKDCDENTCKNCKSEYKMWIDLKSEYEKQV 540
Qy 681 DKYTKDNKKNYDNIDEVNKEANVYLKESKECKOVNFDKIFNESPNEYDMCKKDE 740
Db 541 DKYTKDNKKNYDNIDEVNKEANVYLKESKECKOVNFDKIFNESPNEYDMCKKDE 600
Qy 741 IKYLNIEIKYPTKTHDI 756
Db 601 IKYLNIEIKYPTKTHDI 616

RESULT 17
Q81788_PLAFA PRELIMINARY; PRT; 616 AA.
AC Q81788;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Erythrocyte binding antigen region II (Fragment).
GN Name-eba-140;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22588500; PubMed=12702678;
RA Baum J., Thomas A.W., Conway D.J.;
RT "Evidence for diversifying selection on erythrocyte-binding antigens
of Plasmodium falciparum and P. vivax.";
RL Genetics 163:1327-1336(2003).
DR EMBL; AJ438835; CAD27560.1; -; Genomic_DNA.
FT NON_TER 1
FT NON_TER 1
FT NON_TER 616
SQ SEQUENCE 616 AA; 73526 MW; 311F5009C1744A3F CRC64;

Query Match 51.8%; Score 3360; DB 2; Length 616;
Best Local Similarity 99.4%; Pred. No. 2.7e-133;
Matches 612; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 141 QYTFIQKRTLFPACGIKRSIKWICRENSSEKITVCVPDRKIQLCVANFLNRLTMEKFK 200
Db 1 QYTFIQKRTLFPACGIKRSIKWICRENSSEKITVCVPDRKIQLCVANFLNRLTMEKFK 60
Qy 201 EIFLISVNTAEKLLYNKNEGKPSIFCNELRNSFSDFRSGFIQDDMDFGNTDRVKGYN 260

```

```

Db 61 EIFLISVNTAEKLLYNKNEGKPSIFCNELRNSFSDFRSGFIQDDMDFGNTDRVKGYN 120
Qy 261 TKFSDDYKKEKVKELNNIKKEWMEKNKANLWNHMIWNHKGNIISKECAIIPAEBPQINLWI 320
Db 121 KFSDDYKKEKVKELNNIKKEWMEKNKANLWNHMIWNHKGNIISKECAIIPAEBPQINLWI 180
Qy 321 KEWNEFLMEKRLFLNIKDCKVENKYEACFGGCRLPCCSSYTSFMKKSQTQMEVLTNLY 380
Db 181 KEWNEFLMEKRLFLNIKDCKVENKYEACFGGCRLPCCSSYTSFMKKSQTQMEVLTNLY 240
Qy 381 KKNKSGVDKNNFLNDLFKKNKNKNDLDDFFKNEKEYDDLCDCRYTATTIISKFLNGPAKNDV 440
Db 241 KKNKSGVDKNNFLNDLFKKNKNKNDLDDFFKNEKEYDDLCDCRYTATTIISKFLNGPAKNDV 300
Qy 441 DIASQINVNDLRGFGCNYKSNNEKSNWCTGTFTNKPFGTCEPPTROTLCIGRTYLLHRGH 500
Db 301 DIASQINVNDLRGFGCNYKSNNEKSNWCTGTFTNKPFGTCEPPTROTLCIGRTYLLHRGH 360
Qy 501 EEDYKEHLLGASIEAQLLYKYKEKDENALCSIIQNSYADLADIIGSDIISKDYGGKQM 560
Db 361 EEDYKEHLLGASIEAQLLYKYKEKDENALCSIIQNSYADLADIIGSDIISKDYGGKQM 420
Qy 561 EENLNKVNKDKRNEESLKIIFREKWDENKENVKMSAVLNKNETCKDYDKFKQIPQFL 620
Db 421 EENLNKVNKDKRNEESLKIIFREKWDENKENVKMSAVLNKNETCKDYDKFKQIPQFL 480
Qy 621 RWFKEWGGDFCEKREKIYSFESFKVECKKDCDENTCKNCKSEYKMWIDLKSEYEKQV 680
Db 481 RWFKEWGGDFCEKREKIYSFESFKVECKKDCDENTCKNCKSEYKMWIDLKSEYEKQV 540
Qy 681 DKYTKDNKKNYDNIDEVNKEANVYLKESKECKOVNFDKIFNESPNEYDMCKKDE 740
Db 541 DKYTKDNKKNYDNIDEVNKEANVYLKESKECKOVNFDKIFNESPNEYDMCKKDE 600
Qy 741 IKYLNIEIKYPTKTHDI 756
Db 601 IKYLNIEIKYPTKTHDI 616

RESULT 18
Q81786_PLAFA PRELIMINARY; PRT; 616 AA.
AC Q81786;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Erythrocyte binding antigen region II (Fragment).
GN Name-eba-140;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22588500; PubMed=12702678;
RA Baum J., Thomas A.W., Conway D.J.;
RT "Evidence for diversifying selection on erythrocyte-binding antigens
of Plasmodium falciparum and P. vivax.";
RL Genetics 163:1327-1336(2003).
DR EMBL; AJ438847; CAD27572.1; -; Genomic_DNA.
FT NON_TER 1
FT NON_TER 1
FT NON_TER 616
SQ SEQUENCE 616 AA; 73455 MW; 18DE19BFB9C82582 CRC64;

Query Match 51.8%; Score 3359; DB 2; Length 616;
Best Local Similarity 99.7%; Pred. No. 3e-133;
Matches 614; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 141 QYTFIQKRTLFPACGIKRSIKWICRENSSEKITVCVPDRKIQLCVANFLNRLTMEKFK 200
Db 1 QYTFIQKRTLFPACGIKRSIKWICRENSSEKITVCVPDRKIQLCVANFLNRLTMEKFK 60
Qy 201 EIFLISVNTAEKLLYNKNEGKPSIFCNELRNSFSDFRSGFIQDDMDFGNTDRVKGYN 260

```


Db 61 EIFLISVNTAEKLLYNKNEGDPISFCNELRNSPDSFRSSFIGDMDFGGNTDRVKGYN 120
Qy 261 TKFSDYKKEKNNVEKLNIIKKEWKEKKNANLWNHMIKKECAIIPAEEPQINLWI 320
Db 121 KKFSDYKKEKNNVEKLNIIKKEWKEKKNANLWNHMIKKECAIIPAEEPQINLWI 180
Qy 321 KEWNEFLMEKRLFLNIKDKCVENKKEYEACFGGCRPLPCSSYTSFMKSKTKQMEVLNLY 380
Db 181 KEWNEFLMEKRLFLNIKDKCVENKKEYEACFGGCRPLPCSSYTSFMKSKTKQMEVLNLY 240
Qy 381 KKKNSGVDKNNFLNDLFCNNKNDLDDFFKNEKEYDDLCDCRYTATIIKSPFLNGPAKNDV 440
Db 241 KKKNSGVDKNNFLNDLFCNNKNDLDDFFKNEKEYDDLCDCRYTATIIKSPFLNGPAKNDV 300
Qy 441 DIASQINVDLGRFGCNKSNNEKSWNCTGFTNKPFGTCPPRRQTLCLGRTYLLHRGH 500
Db 301 DIASQINVDLGRFGCNKSNNEKSWNCTGFTNKPFGTCPPRRQTLCLGRTYLLHRGH 360
Qy 501 BEDYKEHLGLGASIIYEAQLKYYKEKDENALCSIIQNSYADLADIIGSDIIKDYGGKQV 560
Db 361 BEDYKEHLGLGASIIYEAQLKYYKEKDENALCSIIQNSYADLADIIGSDIIKDYGGKQV 420
Qy 561 BENLKNVNDKRRNEESLKIIPREKWDENKENVKMSAVLNKNETCKDYDKFKQIPQFL 620
Db 421 BENLKNVNDKRRNEESLKIIPREKWDENKENVKMSAVLNKNETCKDYDKFKQIPQFL 480
Qy 621 RWFKEWGDPCCKEKEKLYSPESFKVCKCKDCDENTCKNCKSKYKWKWIDLKSEYKQV 680
Db 481 RWFKEWGDPCCKEKEKLYSPESFKVCKCKDCDENTCKNCKSKYKWKWIDLKSEYKQV 540
Qy 741 KYLNEIKYPTKXKHI 756
Db 601 KYLNEIKYPTKXKHI 616

RESULT 19

Q66PM7 PLAFPA PRELIMINARY; PRT; 606 AA.
ID Q66PM7 PLAFPA PRELIMINARY; PRT; 606 AA.
AC Q66PM7;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Erythrocyte binding protein-2 (Fragment).
GN Name=EBP-2;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]

NUCLEOTIDE SEQUENCE.
RC STRAIN=ALR, BHZ, FFS, GOS, GVM, IPN, JSL, LANA, NSI, RPN, and 21Q;
RC PubMed=15385490; DOI=10.1128/IAI.72.10.5886-5891.2004;
RA Lobo C.A., de Frazao K., Rodriguez M., Reid M., Zalis M.,
RA Lustigman S.;
RA "Invasion Profiles of Brazilian Field Isolates of Plasmodium
falciparum: Phenotypic and Genotypic Analyses.";
RL Infect. Immun. 72:5886-5891 (2004).
DR EMBL; AY652782; AAT99597.1; -; mRNA.
DR EMBL; AY652783; AAT99598.1; -; mRNA.
DR EMBL; AY652784; AAT99599.1; -; mRNA.
DR EMBL; AY652785; AAT99600.1; -; mRNA.
DR EMBL; AY652786; AAT99601.1; -; mRNA.
DR EMBL; AY652787; AAT99602.1; -; mRNA.
DR EMBL; AY652788; AAT99603.1; -; mRNA.
DR EMBL; AY652789; AAT99604.1; -; mRNA.
DR EMBL; AY652790; AAT99605.1; -; mRNA.
DR EMBL; AY652793; AAT99608.1; -; mRNA.
DR EMBL; AY652780; AAT99595.1; -; mRNA.
FT NON_TER 1

FT NON_TER 606 606
SQ SEQUENCE 606 AA; 72126 MW; 91B73964716D7292 CRC64;
Query Match 51.2%; Score 3319; DB 2; Length 606;
Best Local Similarity 100.0%; Pred. NO. 1.4e-131;
Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 151 LFACIGIKRKSIKICRENSKITTVCVDPDKIQLCVAFNLSRLTMEKFEIFLISVNT 210
Db 1 LFACIGIKRKSIKICRENSKITTVCVDPDKIQLCVAFNLSRLTMEKFEIFLISVNT 60
Qy 211 AKLLYNKNEGKDSIFCNELRNSPDSFRSSFIGDMDFGGNTDRVKGYNITKFSDYK 270
Db 61 AKLLYNKNEGKDSIFCNELRNSPDSFRSSFIGDMDFGGNTDRVKGYNITKFSDYK 120
Qy 271 NVEKLNNIKKEWKEKKNANLWNHMIKKECAIIPAEEPQINLWIKENENFLME 330
Db 121 NVEKLNNIKKEWKEKKNANLWNHMIKKECAIIPAEEPQINLWIKENENFLME 180
Qy 331 KKRLLFNIKDKCVENKKEYEACFGGCRPLPCSSYTSFMKSKTKQMEVLNLYKKNKSGVDKN 390
Db 181 KKRLLFNIKDKCVENKKEYEACFGGCRPLPCSSYTSFMKSKTKQMEVLNLYKKNKSGVDKN 240
Qy 391 NFNLDLFCNNKNDLDDFFKNEKEYDDLCDCRYTATIIKSPFLNGPAKNDVIAQINVD 450
Db 241 NFNLDLFCNNKNDLDDFFKNEKEYDDLCDCRYTATIIKSPFLNGPAKNDVIAQINVD 300
Qy 451 LRFGGCKYKSNNEKSWNCTGFTNKPFGTCPPRRQTLCLGRTYLLHRGHEEDYKEHLG 510
Db 301 LRFGGCKYKSNNEKSWNCTGFTNKPFGTCPPRRQTLCLGRTYLLHRGHEEDYKEHLG 360
Qy 511 ASIYEAQLKYYKEKDENALCSIIQNSYADLADIIGSDIIKDYGGKMEENLNKVKND 570
Db 361 ASIYEAQLKYYKEKDENALCSIIQNSYADLADIIGSDIIKDYGGKMEENLNKVKND 420
Qy 571 KKRNEESLKIIPREKWDENKENVKMSAVLNKNETCKDYDKFKQIPQLRWFKEWGD 630
Db 421 KKRNEESLKIIPREKWDENKENVKMSAVLNKNETCKDYDKFKQIPQLRWFKEWGD 480
Qy 631 CEKREKIIYSPESFKVCKCKDCDENTCKNCKSKYKWKWIDLKSEYKQVDTKDKKK 690
Db 481 CEKREKIIYSPESFKVCKCKDCDENTCKNCKSKYKWKWIDLKSEYKQVDTKDKKK 540
Qy 691 MYDNIDEVKNEANVYLKESKECKVNFDDKIFNESPNEVEDMCKKDEIKYLNKYP 750
Db 541 MYDNIDEVKNEANVYLKESKECKVNFDDKIFNESPNEVEDMCKKDEIKYLNKYP 600
Qy 751 KTKXHI 756
Db 601 KTKXHI 606

RESULT 20

Q66PL8 PLAFPA PRELIMINARY; PRT; 606 AA.
ID Q66PL8 PLAFPA PRELIMINARY; PRT; 606 AA.
AC Q66PL8;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Erythrocyte binding protein-2 (Fragment).
GN Name=EBP-2;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]

NUCLEOTIDE SEQUENCE.

RC STRAIN=04Q, PSS1, and 350;
RC PubMed=15385490; DOI=10.1128/IAI.72.10.5886-5891.2004;
RA Lobo C.A., de Frazao K., Rodriguez M., Reid M., Zalis M.,
RA Lustigman S.;
RA "Invasion Profiles of Brazilian Field Isolates of Plasmodium
falciparum: Phenotypic and Genotypic Analyses.";
RL Infect. Immun. 72:5886-5891 (2004).

```
DR EMBL; AY652791; AAT99606.1; -; mRNA.
DR EMBL; AY652792; AAT99607.1; -; mRNA.
DR EMBL; AY652791; AAT99596.1; -; mRNA.
FT NON_TER 1
FT NON_TER 606
SQ SEQUENCE 606 AA; 72194 MW; 85526E7D089FB22C CRC64;

Query Match 51.1%; Score 3309; DB 2; Length 606;
Best Local Similarity 99.5%; Pred. No. 3.6e-131;
Matches 603; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 151 LFACGIRKRSIKWICRENSSEKITVCVPDRKIQLCVANFLNSRLTMEKFEIFLISVNTS 210
DB 1 LFACGIRKRSIKWICRENSSEKITVCVPDRKIQLCVANFLNSRLTMEKFEIFLISVNTS 60

QY 211 AKLLYNKEGKPSIFCNELRNSFSDRSFSGFIDGMDFGGNTDRVKGYINTKFSDDYK 270
DB 61 AKLLYNKEGKPSIFCNELRNSFSDRSFSGFIDGMDFGGNTDRVKGYINTKFSDDYK 120

QY 271 NVEKLNNIKKEWWEKNKANLWNHMIVNHKGNISKECAIIPAEBPQINLWIKENNFME 330
DB 121 NVEKLNNIKKEWWEKNKANLWNHMIVNHKGNISKECAIIPAEBPQINLWIKENNFME 180

QY 331 KKRFLNIDKCVENKYEACFGGCRPCSSYTSFMKSKTOMEVLTNLKKNKSGVDKN 390
DB 181 KKRFLNIDKCVENKYEACFGGCRPCSSYTSFMKSKTOMEVLTNLKKNKSGVDKN 240

QY 391 NFLNDLFKKNKNDLDDFFPKNEKEYDDLCDCRYTATTIISKFLNGPAKNDVDIASQINVD 450
DB 241 NFLNDLFKKNKNDLDDFFPKNEKEYDDLCDCRYTATTIISKFLNGPAKNDVDIASQINVD 300

QY 451 LRFGGNYKNNEKSNWCTGTFTNKPFGTCPPRRQTLCLGRVYLLHRGHEEDYKEHLG 510
DB 301 LRFGGNYKNNEKSNWCTGTFTNKPFGTCPPRRQTLCLGRVYLLHRGHEEDYKEHLG 360

QY 511 ASIYEAOQLLYKYEKEDENALCSIIQNSYADLADIIGKSDIISKDYGKMEENLNKVKD 570
DB 361 ASIYEAOQLLYKYEKEDENALCSIIQNSYADLADIIGKSDIISKDYGKMEENLNKVKD 420

QY 571 KGRNEESLKIIFREKWDENKENVKMSAVLNKNETCKDYDKFQKIPOFLRWFKEWGDDF 630
DB 421 KGRNEESLKIIFREKWDENKENVKMSAVLNKNETCKDYDKFQKIPOFLRWFKEWGDDF 480

QY 631 CERKEKIIYFSESPKVECKKDCDENTCKNCKSEYKWKIIDLKSEYKQVDYTKDKNKK 690
DB 481 CERKEKIIYFSESPKVECKKDCDENTCKNCKSEYKWKIIDLKSEYKQVDYTKDKNKK 540

QY 691 MYDNIDVKNKEANVYLKESKECKDVNFDDKIFNESPNYEYDMCKKDEIKYLNEIKYP 750
DB 541 MYDNIDVKNKEANVYLKESKECKDVNFDDKIFNESPNYEYDMCKKDEIKYLNEIKYP 600

QY 751 KTKHDI 756
DB 601 KTKHDI 606

RESULT 21
Q81760 PLARE PRELIMINARY; PRT; 616 AA.
AC Q81760;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Putative erythrocyte binding antigen region II (fragment).
GN Name-eba-140 homologue;
OS Plasmodium reichenowi;
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5854;
RN [1]
RX NUCLEOTIDE SEQUENCE.
RA MEDLINE=22588500; PubMed=12702678;
RA Baum J., Thomas A.W., Conway D.J.;
RT "Evidence for diversifying selection on erythrocyte-binding antigens

RT of Plasmodium falciparum and P. vivax.";
RL Genetics 163:1327-1336(2003).
DR EMBL; AJ438829; CAD27554.1; -; Genomic_DNA.
FT NON_TER 1
FT NON_TER 616
SQ SEQUENCE 616 AA; 73518 MW; DAA91E321E6110E1 CRC64;

Query Match 48.5%; Score 3144; DB 2; Length 616;
Best Local Similarity 92.2%; Pred. No. 3.1e-124;
Matches 568; Conservative 20; Mismatches 28; Indels 0; Gaps 0;

QY 141 QYTFIQKRTHLFACGIRKRSIKWICRENSSEKITVCVPDRKIQLCVANFLNSRLTMEKFK 200
DB 1 QYTFIQKRTHLFACGIRKRSIKWICRENSSEKITVCVPDRKIQLCVANFLNSRLTMEKFK 60

QY 201 EIFLISVNTAEKLLYNKNEGKPSIFCNELRNSFSDRSFSGFIDGMDFGGNTDRVKGYIN 260
DB 61 EIFLISVNTAEKLLYNKNEGKPSIFCNELRNSFSDRSFSGFIDGMDFGGNTDRVKGYIN 120

QY 261 TKFSDYDYKKNVEKLNIIKKEWWEKNKANLWNHMIVNHKGNISKECAIIPAEBPQINLWI 320
DB 121 MKFSDYDYKKNVEKLNIIKKEWWEKNKANLWNHMIVNHKGNISKECATIPEEBPQINAWI 180

QY 321 KENNFNPLMEKRLFLNIDKCVENKYEACFGGCRPCSSYTSFMKSKTOMEVLTNL 380
DB 181 KENNFNPLMEKRLFLNIDKCVENKYEACFGGCRPCSSYTSFMKSKTOMEVLTNL 240

QY 381 KKNKSGVDKNNFLNDLFKKNKNDLDDFFPKNEKEYDDLCDCRYTATTIISKFLNGPAKNDV 440
DB 241 KKNKSGVDKNNFLNDLFKKNKNDLDDFFPKNEKEYDDLCDCRYTATTIISKFLNGPAKNDV 300

QY 441 DIASQINVDLRFGGNYKNNEKSNWCTGTFTNKPFGTCPPRRQTLCLGRVYLLHRGH 500
DB 301 DTASKINVDLRVFGGNYKNNEKSNWCTGTFTNKPFGTCPPRRQTLCLGRVYLLHRGH 360

QY 501 EEDYKEHLGASIEAQLLYKYEKEDENALCSIIQNSYADLADIIGKSDIISKDYGK 560
DB 361 EEDYKEHLGASIEAQLLYKYEKEDENALCSIIQNSYADLADIIGKSDIISKDYGK 420

QY 561 EENLNKVKNDKGRNEESLKIIFREKWDENKENVKMSAVLNKNETCKDYDKFQKIPOFL 620
DB 421 EESLNKVKNDKGRNEESLKIIFREKWDENKENVKMSAVLNKNETCKDYDKFQKIPOFL 480

QY 621 RWPKEWGDDFFCEKREKIIYFSESPKVECKKDCDENTCKNCKSEYKWKIIDLKSEYKQV 680
DB 481 RWPKEWGDDFFCEKREKIIYFSESPKVECKKDCDENTCKNCKSEYKWKIIDLKSEYKQV 540

QY 681 DKYTKDKNKKVNDIDVKNKEANVYLKESKECKDVNFDDKIFNESPNYEYDMCKKDE 740
DB 541 ERYTKDKNNIYDNIDVKNKEANVYLKESKECKDVNFDDKIFNESPNYEYDMCKKDE 600

QY 741 IKYLNKIKYKTKHDI 756
DB 601 IKYLNKIKYKTKHDI 616

RESULT 22
Q25842 PLAPA PRELIMINARY; PRT; 1475 AA.
AC Q25842;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Erythrocyte binding protein.
GN Name-erythrocyte-binding protein;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RX NUCLEOTIDE SEQUENCE.
RA MEDLINE=92357776; PubMed=1496004;
RA Adams J.H., Sim B.K., Dolan S.A., Fang X., Kaslow D.C., Miller L.H.;
RT "A family of erythrocyte binding proteins of malaria parasites."
```

Proc. Natl. Acad. Sci. U.S.A. 89:7085-7089 (1992).	753	---KHIYDID---	TFSDTFGDFGPISI---	774	
EMBL; M93397; AAA29600.1; -; Genomic DNA.	865	RPDKKELEDQNSDESEETVNVHISKSPSINNGDDSGSATVSESSNSTGLSIDDDRNG	924		
GO: 0016021; C: integral to membrane; IEA.	775	-----NANINEQSGKDTSTNGSETSDSP---	VSHEPESDAINV	812	
GO: 0004872; F: receptor activity; IEA.	925	DTFVETQTANTVEDVIRKENADKDEKGADEERKSTSESLSPEEKMLTDNEGNSLNH	984		
GO: 0009405; P: pathogenesis; IEA.	813	EKL-----SGDESSSETRGILDIN-----	DPSVTNNVNEVH---	846	
InterPro: IPR008602; Duffy binding.	985	BEVKEHTNSNDVQSGGIVNMVVEKELKOTLENPSSLDDEKGAHELSBPNLSLSDQDMS	1044		
Pfam: PF05424; Duffy binding, 2.	847	NTQGSVSTSDITN---CHSESSLNR-----	TTNAQDIKIRSGNEQSDN---	OENS	892
SEQUENCE 1475 AA; 171487 MW; EB8312823APF946D CRC64;	1045	NTPGPLDSTETTERISNNKEYKVEREDERTLTKEYEDIVLAKSHMNRSDDGELYDENS	1104		
	893	SHSSNSGSLIT-GOVPSDNTQNTY-----	DSNPHRDTPNALA-----	SLP-	934
	1105	DUSTVNDSEDAEAKMGKNDTSEMSHNSQHIESDQKNDKMTVGDLTGTHVONEISVPV	1164		
	935	---SDKINE-----	IEGFDSDSDSNGRDTTSTNTHDVR-----	R	967
	1165	TGEIDKLRSEKSIKAEERLSHTDIHKINPEDRNSNTLHLKOIRNEENERHLTNQ	1224		
	968	TNIVSERVRVSHDFIRNGMANNNAHQVITQ-----	IENNGIIRGOEESAGNSVNYKDN-	1021	
	1225	INISQERDLQKHGF-----	HTMNNLHGDGVSERSQINSHSHHGNQDRGSGNSGNLNMRSNN	1280	
	1022	-----PKRSNPFSSNDHKNIQEVNSRDTKRVPEETIKLSKONKCNNEYSMEYCTYSDE	1075		
	1281	NNFNIPRYNL-----	YDKKLDLDLYENRNDSTTKELIKKLAIEKNCENELSVKYCDHMIH	1337	
	1076	RNSSPGCSREBERKKLCCQISDYCLYFNFYBYNYNCISEKSPRYCKFKSEGOSSIP	1135		
	1338	BEIPLKCTCKETRNLCQAVSDYCYFTYDSEYNYNCTKREPDSPYTCFKEAFSSMP	1397		
	1136	YPAAGILVIVLLSSASRMGKSNSEY--DIGESNTEATPEENNYLNKLSRIFNQEVOE	1193		
	1398	YVAGAGLVFIILVILGASQAKYQSSEGVMNENNENF--LFEVTDNLDKLSNMFNQVOE	1455		
	1194	TNISDYSEYN 1203			
	1456	TNINDFSEYH 1465			
	RESULT 23				
	Q81BE8_PLAF7				
	ID Q81BE8_PLAF7 PRELIMINARY;	PRT;	1462 AA.		
	AC Q81BE8;				
	DT 01-MAR-2003 (TRENBLrel. 23, Created)				
	DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)				
	DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)				
	DE Erythrocyte binding antigen.				
	GN Name=eba-175; Synonyms=MAJ7P1.176, PF07_0128;				
	OS Plasmodium falciparum (isolate 3D7).				
	OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.				
	OX NCBI_TaxID=36329;				
	RN [1]_TaxID=36329;				
	RP NUCLEOTIDE SEQUENCE.				
	RA Seeger K., Murphy L., Harris D., Berriman M., Pain A., Hall N.,				
	RA Quail M., Barrall B.;				
	RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.				
	DR EMBL; AL844506; CAD51055.1; -; Genomic DNA.				
	DR GO; GO:0016021; C: integral to membrane; IEA.				
	DR GO; GO:0004872; F: receptor activity; IEA.				
	DR GO; GO:0009405; P: pathogenesis; IEA.				
	DR InterPro; IPR008602; Duffy binding.				
	DR Pfam; PF05424; Duffy binding, 2.				
	SQ SEQUENCE 1462 AA; 169909 MW; 412B875061E60AA5 CRC64;				
	Query Match	24.6%;	Score 1596.5;	DB 2;	Length 1462;
	Best Local Similarity	28.2%;	Pred. No. 7.2e-59;		
	Matches 425;	Conservative 221;	Mismatches 476;	Indels 385;	Gaps 46;

Db	925	DTFVRTQDTANTEDVIRKENADKDEKGADEBRHSTSESLSPEEKMLTDNEGNSLNH	984
Qy	813	EKL---SGDESSSETGILDIN-----DPSVTNNVNEVH-----DAS	846
Db	985	EEVKEHTSNDNVQOOSGIIVNMVVEKELKOTLENPFSSSLDEGKAHELSSEPNLSSDOOMS	1044
Qy	847	NTOGSVNSTDITN--GHSESSLNR-----TTNAODIKIGRSGNEQSDN-----QENS	892
Db	1045	NTEPLDNTSEETTERISNNYKVNEREDERTLTKEYEDVLKSHMNRSEDGDGLYDENS	1104
Qy	893	SHSDNSGSITI-QGVPSIDENTQWY-----DSQNPRDTPNALA-----SLP-	934
Db	1105	DLSTVNDESDEARAKMKGNDDTSEMHSNQIHESDOOKDMKTVDGLGTTHVQNEISVPV	1164
Qy	935	---SDDKINE-----LEGFDSRDSENGRGDTTSNTHDVR-----R	967
Db	1165	TGEIDEKLRESKSKTHKAEERLSHTDIHKINPEDRNSNLHLKDIREENERHLTQN	1224
Qy	968	TNIVSERRVNSHPFRINGMANNAHHQIYTQ-----IENNGIIRGOBESAGNSVNYKDN-	1021
Db	1225	INISOERDLQKHGF----HTWMNLHGDSVRSQSINHSHGRNQDRCGNSGNVLNMRSN	1280
Qy	1022	-----PKRSNFSENDDKKNIQSYNSRDKTVREBIILKSQNKCNWEYSMEYCTYSDE	1075
Db	1281	NNFNNPSPRYNL---YDKLDLDLYENRNDSTTKELIKKLARIKNCEISVKYCDHMII	1337
Qy	1076	RNSPGPCSRERKCLCOISDYCLKYFNFSYTYNCIKSEIKSPKYCFKSEGQSSI	1134
Db	1338	EBIPLKTCRKTRNLCCAVIDCMSYFTYDSEYNYCTKREFDDPSYTCRKEAFSSM	1396
RESULT 25			
Q9NG63_PLAFA			
ID	Q9NG63_PLAFA PRELIMINARY;	PRT; 1435 AA.	
AC	Q9NG63;		
DT	01-OCT-2000 (TrEMBLrel. 15, Created)		
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE	Erythrocyte binding antigen 175.		
OS	Plasmodium falciparum.		
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
OX	NCBI_TaxID=5833;		
RN	[1]		
RA	NUCLEOTIDE SEQUENCE.		
RL	Ma C.L., Yu X.B., Li X.R., Shan Z.X.; Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF258781; AAF2186.1; -, Genomic DNA.		
DR	GO; GO:0016021; C:integral to membrane; IEA.		
DR	GO; GO:0004872; F:receptor activity; IEA.		
DR	GO; GO:0009405; P:pathogenesis; IEA.		
DR	InterPro; IPR008602; Duffy binding.		
DR	Pfam; PF05424; Duffy_binding_2.		
SQ	SEQUENCE 1435 AA; 167269 MW; FP958828C7794F22 CRC64;		
Query Match			
Best Local Similarity 28.4%; Pred. No. 8.4e-57;			
Matches 410; Conservative 216; Mismatches 460; Indels 358; Gaps 44;			
Qy	1	MKGYNFYFLPIFLIFYLVNVRINESIIGRTLNYNRQDESDISRVSNGPELANNHKTNIYDS	60
Db	1	MKCNISYIFPASFPVLVFYAKARNEYDI-----KENEFKFLDVYKEKFELDKKKYGNVQKT	55
Qy	61	DYEDVNNKLINSVENKSVKCKBSLFINN-----KTSKY---DIIPPSYSYENDKFNSL	112
Db	56	D-----KKIFTFIENK-----LDLANSFKFNKWSYGTDPNDIKNNLSLIN-----	96
Qy	113	SENEDSGNTSNNFANTSISIGKNKQ-----YTIQKRTH-----	150
Db	97	KHNNEFMFNNOYSFLSTS--SLIKQKYVINAVRVSRILSLFDSEINNGRTSSNNEV	154
Qy	151	LFAQGIGKRSIKWICRENSEKIT-VCPVPRKIQLCVANFLNSRLETMEKPKFIPLSVNT	209

Db 155 LSNCREKRGKWKDKKKNDKRSNTVCIPDRRIQICVNLISIIKTYTETWTDHFIASKK 214
Qy 210 EAKLLYNKNEGKPSIFCNELRNSFSDFRSSFGDDMDFGGNTDRVKGYNITKFSDYIKE 269
Db 215 ESQLLKKNDKNYNSKFCNDLKNISFLDYGHLAGNDMDFGYSTKAENKIQEVFKGAHGK 274
Qy 270 KNYEKLANNIKKEWKKANLANLWMMIYNHKNISKECAIIPABEPQINLWIKENENFLM 329
Db 275 ISEHEIKFRKKNWNEFREKLWEAMLSEHKNIN-NCKNIPQEBELQITQWIKWHGFEFL 333
Qy 330 EKRLFLNIDKCKVENKYEACFGGCRLLPCSSYTSFMKSKTOMELVNLTKKNSGVDK 389
Db 334 ERYNRSKLPKSKCKNNTLYEACEKECIDPCMKYRDMIRSKFPMHTLSKEYETQN--VSK 391
Qy 390 NNFLNDLFLK-KNNKND--LDDFFKN-EKEYDDLCDCRYTATIIKSFNLGPAKNDVDIASQ 445
Db 392 ENAENYLIKISKNKNDKAKVSLNLCDAEYSKYCDCKHHTTLVKSVLNGNDNTIKEREH 451
Qy 446 INVNDLRGFCNYKS--NNEKSNWCTGTFNKFPGTCPPRRQTLCLGRYTLHRGHEED 503
Db 452 IDLDDFKGFCGDKNSVDNTNTKWECKPYILSTKDVCPVPPRRQELCLGNIDRIYDKNLLM 511
Qy 504 YKHELLGASIEAQLLYKVKEDENALCSTIQNSYADLADIKGSDIIDYVYKGMKEEN 563
Db 512 IKEHILAIAYESILKRRKYNKDDKEVKLIINKTFADIRDIIGTDYWNLDLSNRKLVGK 571
Qy 564 LNKVNKDKRNEESLKIIFREKWNDEKENYKVMASVAVLNKNETKOYDFKQKIPQFLRW 623
Db 572 INTNSYVHRKNDKLFREDEWVKIKDVVNVISWFKDKVCKE--DDIENIPQPPRWF 630
Qy 624 KENGDDCEKKEKIEYFESFKVECKKDCDENTCKKCSYKWKIDLKSEYKQV--- 680
Db 631 SEWGGDYCDQKTKMI---ETLKVECKEPCEDDNCCKNSKYEWISKEEYENKQAKQY 687
Qy 681 DKYTKDNKKMYDNIIDVKNKEANVLYKESKECKDVNDDKIFNESPNEYDMCKKDE 740
Db 688 QEYQKGNYYMY---SEFKSIKPEVLYKYSEKSNLNFDEDFKELHSDYKKNCTWCPE 744
Qy 741 IKYL-----NB----- 746
Db 745 VKDVPISIRNEQTSQEAPEENTEITHRTETPSISEGPKGNEQKERRDDSLSKISVSP 804
Qy 747 -----IKPKT----- 752
Db 805 ENSRPETDAKDTSNLLKLGVDISMPKAVIGSPNDNINVTQGDNISGVNSKPLSDDV 864
Qy 753 ---KHDYVID---TFSDTFGDTGPISI----- 774
Db 865 RPKKELEDQNSDESEETVNVHISKSPSINNGDSSGSGSATVSSSSNTGLSIDDDRNG 924
Qy 775 -----NANINEQOSGKDTNTGNSSETSDP---VSHEPESDAAINV 812
Db 925 DTFVRTQDTANTEDVIRKENADKDEKGADEERHSTSESLSPPEKMWSTNEGGSNLH 984
Qy 813 EKL-----SGDESSSETGILDIN-----DPSTVNAVNEVH----- 846
Db 985 EEVKEHASNDNVQOSGGINVMNYVEKELDNTLENPSSSLDEGKAHELSFNLSSDQMS 1044
Qy 847 NTQGSVNTSDITN---GHSESSLNR-----TTNAQDIIKIGRSGNEQSDN---QENS 892
Db 1045 NTPGLDNTSEETTERISNNEYKYNREGERLTLYKEYEDIVLKSHMRESDDGELYDENS 1104
Qy 893 SHSSDNGSLTI-QQVPESDNTQNTY-----DSQNPHRDTPNALA-----SLP- 934
Db 1105 DLSTVNDESEDAEAKMGNDTSEMHSNOSHIESDQKNDMKTGDLGTHVQNEISVPV 1164
Qy 935 ---SDDKINE-----LEGDFSRRDSENGGDTTSNTHDVR-----R 967
Db 1165 TGEIDEKLRESKESKIHABEERLSHTDIIKHINPEDRNSNTLHLKDIRNEBERHLTQN 1224
Qy 968 TNIVSERRVNSHDF--JRN-----GWAN-----NNAHQVYITQIENNNGIIRGOESAGNSVN 1017
Db 1225 INTSQERDLQHGPHGTVNNLHGDGVSERSQINSHH-----GNRQDGGNSGVNLN 1275

Qy 1018 YKDN-----PKRSNFSSENHKKNIQEYNSRDTKRVREBIIKLSKQKNNNEYSMEYC 1070
Db 1276 MRSNNNNFNPIPERYNL---YDKKLDSDLYENRNDSTTKRELKLAELKNEISVKYC 1332
Qy 1071 TYSDESNSSPGCPSRERKLCQOISDYCLKYNFYNSIYNYNCIKSSEIKSPYKCPKSEG 1130
Db 1333 DHMIHEEIPLKTCETKTRNLCAVSDYCMYSFTYDSEYNYNCTKREDFDPSYTCRKEA 1392
Qy 1131 QSSI 1134
Db 1393 FSSM 1396

RESULT 26
Q05644_PLAPA
ID Q05644_PLAPA PRELIMINARY; PRT; 1421 AA.
AC Q05644;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Erythrocyte-binding antigen-175 (Fragment).
GN Name=EBA-175;
OS Plasmodium falciparum.
OC Bukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3D7;
RA Daugherty J.R., Lanar D.E.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U32207; AAA75179.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy-binding.
DR Pfam; PF05424; Duffy-binding; 2.
FT NON TER 1421
SQ SEQUENCE 1421 AA; 165330 MW; 7BE8DDFA07CBE771 CRC64;

Query Match 23.7%; Score 1537.5; DB 2; Length 1421;
Best Local Similarity 28.1%; Pred.No. 2.1e-56;
Matches 411; Conservative 211; Mismatches 467; Indels 375; Gaps 44;
Qy 1 MKGYFIYLIPLIFLYNVIRINESIIGRTLYNRQDESSDSRVNSPELNNHKTNIYDS 60
Db 1 MKCNISYIFPASPFLVYFAKARNEYDI-----KENEKFLDVYKEKFNELDKKYGVNQKT 55
Qy 61 DYEDVNNKLINSVENKSVKKGKSLSFINN-----KTKSY---DIIPPSYSYRNDKFNSL 112
Db 56 D-----KKIFTPIENK-----LDLNNSEKFNKWKSYGTDPNIDKNMSLIN----- 96
Qy 113 SENEDSGNTNSNPFANTSEISIGDKNQ-----YTFIQKETH----- 150
Db 97 KHNEEMFNNNYOSFLUSTS--SLIKQNKYVPINAVRSRILSDSRINNGRNTSSNEV 154
Qy 151 LFAAGIKRKSIIKWCIRENSEKIT-VCVPDRKIQLCVANFLNSLETWKEFKEIFLISVNT 209
Db 155 LSNCREKRGKWKDKKKNDKRSNYVCIPDRRIQICVNLISIIKTYTETWTDHFIASKK 214
Qy 210 EAKLLYNKNEGKPSIFCNELRNSFSDFRSSFGDDMDFGGNTDRVKGYNITKFSDYIKE 269
Db 215 ESQLLKKNDKNYNSKFCNDLKNISFLDYGHLAGNDMDFGYSTKAENKIQEVFKGAHGK 274
Qy 270 KNYEKLANNIKKEWKKANLANLWMMIYNHKNISKECAIIPABEPQINLWIKENENFLM 329
Db 275 ISEHEIKFRKKNWNEFREKLWEAMLSEHKNIN-NCKNIPQEBELQITQWIKWHGFEFL 333
Qy 330 EKRLFLNIDKCKVENKYEACFGGCRLLPCSSYTSFMKSKTOMELVNLTKKNSGVDK 389
Db 334 ERYNRSKLPKSKCKNNTLYEACEKECIDPCMKYRDMIRSKFPMHTLSKEYETQN--VSK 391
Qy 390 NNFLNDLFLK-KNNKND--LDDFFKN-EKEYDDLCDCRYTATIIKSFNLGPAKNDVDIASQ 445

```
Db 392 ENAENYLIKISGNKNDKAVSLLNANCAEYKSKYCDKHTTTLVKSVLNGNDNTIKKREH 451
Qy 446 INVNDLRGFCGNYKS--NNEKSWNCTGFTNKPFGCTEPPROTCLGRTYLLHRGHEED 503
Db 452 IDLDDFSKFGCKNSVDNTNKTWECKPKYKLSKTDVCPVPPRQBLCLGNIDRIIRKLLM 511
Qy 504 YKHLGASIEYEAOLLYKYEKDBNALCSIIIONSADLADIIGSDIIKDYGKMBEEN 563
Db 512 IKEHLAIAIYESRLKRYKNDKKEVCKIINKTFADIRIIGTDYNDLSNRKLVGK 571
Qy 564 LNKVNKKRNEBSLKIPREKWDENKENVKMSAVLNKNETCKDYDKPKQIPQPLRWF 623
Db 572 INTNSYVHRNKQNDKLFDESWKVIKDVWVSWFVKDVTCKE--DDIENIPOPFRWF 630
Qy 624 KEMGDDPCKEKEKIEKYSFESKVECKKDCDENTCKNCKSYKWKWIDLKSEYKQV--- 680
Db 631 SEWGDYQDQTKMI---ETLKVCEKPKPCEDNCKRCKNSYKEMI SKKGEYNKQAKQY 687
Qy 681 DKYTKDKKKMYDNIDEVKNEANVYLKKEKCKDVNFDKIFNESNEYEDMCKKDE 740
Db 688 QEQKGNMYK---SEPKSIKPEVYLKYSKCSNLNFEDEFKELHSDYFNKCTWCP 744
Qy 741 IKYL-----SEPKSIKPEVYLKYSKCSNLNFEDEFKELHSDYFNKCTWCP 750
Db 745 VKDVPISIRNEQTSQAVPEBSTEIAHRTETRTDERKNOEPANKOLKNPQQSVGENGT 804
Qy 751 -----KYKHD 755
Db 805 KDLAQEDLGGSRSEDEVTEQFVGNHIGPKGBDOTLKGSDAIPNIGEPETGISTEESRHE 864
Qy 756 -----IYDITSDTF-----GGTPISTINANIEQSG-KDTSNT----- 790
Db 865 EGRNKQALSTSVDPPELSDTLQLHEDTKENDKLFLESSTITPTESGSDTEETPSISEG 924
Qy 791 --GN-----SETSDSPVSHPEPSDA--AINVEKLSGD----- 818
Db 925 PKGNEQKRRDDSLSKISVSPENSREPETAQDNTNLLKLGKDVDSMPKAVIGSSPNDNI 984
Qy 819 -----ESSETGILDIN----- 831
Db 985 NVTEQGDNI SGVNSKPLSDVDPKQNHBEVKEHTSNSDNVQSGGIVNMVYKELKDTLE 1044
Qy 832 DPSVTNNVNEH-----DASNTQGSVNTSDITN--GHSESLNR----- 869
Db 1045 NPSSLDEGKAHELSEPNLSSDDQMSNTPLDNTSEETTERISNNEKYKNERGERTL 1104
Qy 870 TTNAQDIKIGRSGNEQSDN---QENSHSDNSGSLTI-QQVPSDNTQWY----- 917
Db 1105 TKEYEDIVLKSHMRSDDGLYDENSLSVTNDESEDAEAKMGNDTSEMSHNSQHIE 1164
Qy 918 -DSQNHRTDPNALA-----SLP-----SDDKINE-----IEGFDSDRDS 951
Db 1165 SDQCKNDMKTGDLGTHVQNEISVPVTEIDEKLRESKESIKHABEERLSHTDHIKN 1224
Qy 952 ENGRGDTTSNTHYR-----RTNIVSERVNSHDFIRGMANNNAHQYITQ-- 998
Db 1225 PEDRNSNTLHLKORNEENERHLTNQINISQERDLQKHG-----HTWNNLHGQVRS 1280
Qy 999 ---IENNGIIRGQESAGNSVNYKDN-----PKRSNFSSENDHKQNIQYNSRDTKRV 1048
Db 1281 QINSHHGNRQDRGNSGVLNMRNNNNFNINISRYNL---YDKGLDLIYENRDSST 1337
Qy 1049 REBIILKSKONKNNEYSMBYCTYSDESNSSPGCSREERKLCQISDYCLKYPNFYSI 1108
Db 1338 KELIKLABINKCENESVYKCDHMIHEEIPLKTCTKEKTRNLCAVSDYCMSTFTYDSE 1397
Qy 1109 EYVNCIKSEIKSPYKCFKSGQS 1132
Db 1398 EYVNCIKREFDPSYTCFRKAFA 1421
```

RESULT 27

```
Q95VT1_PLAFA
ID Q95VT1_PLAFA PRELIMINARY; PRT; 1402 AA.
AC Q95VT1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Erythrocyte-binding antigen-175 (Fragment).
GN Name=EBA-175;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5933;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Li X.R., Chishti A.H., Oh S.S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF06762; AAK96216.1; -, mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy binding.
DR Pfam; PF05424; Duffy_binding; 2.
FT NON_TER 1
FT NON_TER 1402
SQ SEQUENCE 1402 AA; 163080 MW; A9E21BD8774E429A CRC64;

Query Match 23.5%; Score 1524; DB 2; Length 1402;
Best Local Similarity 28.2%; Pred. No. 7.6e-56;
Matches 402; Conservative 207; Mismatches 446; Indels 370; Gaps 43;

Qy 40 DISRVNSPELNHNKNTIYDSYEDVNNKLINSFVENSKVKKSLSPINN-----KTKS 94
Db 16 DVYKEKFNELDKKYGNNVQKTD-----KKIPTIENK-----LDILNNSKFNKRWKS 62
Qy 95 Y---DIIPPSVSYRNDKFNLSNEDNSGNTNSNPNANTSEISIGKONKQ----- 141
Db 63 YGTPDNIDKMSLIN-----KHNEEMFNNNYQSFLSTS--SLIKONKYVPINAVRSR 114
Qy 142 -YTFIQKRTH-----LFCAGIKRKSIIKWCIRENSEKIT--VCVPDRKIQLCVANF 188
Db 115 ILSPLDRIRNGRNTSSNNEVLSNCRCKRGMKWDCKKNDKRSNYVICPDRIQLCIVNL 174
Qy 189 LNSRLETWKEFKEIFLISVNTAEKLYNKNEGDKPSIFCNELRNSFSDRFSFGDDMDF 248
Db 175 SLIKTYTKETKOHFIEASKESQLLLKNDKYNKSKFCDLKNFLDYGHLAGNDMDF 234
Qy 249 GGNTDRVKGIVNTKPSDYKKEKVEKLNNTIKGEWEKKNKANLMNHMIVNHKGNISKECAI 308
Db 235 GGYSTKAENKIQVFKGAHGEISEHKIKNPKRKEWNEFREKLEWAMLSEHKNNIN--NCKN 293
Qy 309 IPAREPQINLIKEMWENFLWEKRLFNLIKDKCVENKYEACFGGCLPCSSYTSFPMKK 368
Db 294 IPOBELQITQIKWEHGEFLERDRNRSKLPKSKNNNTLYEACEKICIDPCMKYRDMIR 353
Qy 369 SKTQMEVLTMLYKKKNSGVDKNNFLNLFK--KNNKD--LDDFFKN--EKEYDDLCDCRYT 424
Db 354 SKFEWHTLSKEYETOK--VPKENAEVYLIKISEKNDAKVSLLNNDCAEYSKYCDCKHT 411
Qy 425 ATTIKSLFNGPAKNDVDIASQINVDLRGFCGNYKS--NNEKSWNCTGFTNKPFGCTEPP 482
Db 412 TTLVKSVLNGNDNTIKEREHIDLDLDFSKFGCDKNSVDNTNKTWECKPKYKLSKTDVCP 471
Qy 483 PRQOTLCIGRTYLLHRGHEEDYKHLGASIEYEAOLLYKYEKDBNALCSIIIONSADL 542
Db 472 PRQELCLGNIDRIYDKNLLMIKEHLAIAIYESRLKRYKNDKKEVCKIINKTFADI 531
Qy 543 ADIIKGSDIIKDYGKMBEENLANKVKDKKRNEBSLKIPREKWDENKENVKMSAVLNK 602
Db 532 RDIIGGTDYNDLSNRKLVGKINTNSYVHRNKQNDKLFDESWKVIKDVWVSWFVK 591
Qy 603 NKETCKDYDKPKQIPQFLRWPKWGDDDFCERKREKIIYSFBSFKVECKKQDCDENTCK 662
Db 592 DKTCKE--DDIENIPOPFRWFSEWGDYQDQTKMI---ETLKVCEKPKPCEDNCKRKC 647
```



```
QY 663 SEYKKNWDLKSEYKQV---DKYTKDNKKNMYNDIVKKNKEANVYLKSKSECKDVNF 719
D  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 648 NSYKEMISKKBEYNKQAKQYQYQKGNVY---SEFKSIKPEVYLKYSEKSNLNF 704
QY 720 DDKIFNENPNYEDMKCKDBEIKYL----- 744
D  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 705 EDEPKELHSDYKKNCTWCPKVDKVPISIRNNEQTSQEAVPBESIEAHRTTRDERK 764
QY 745 -----NEIKYP----- 750
D  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 765 NOEPANKDLNPOQSVGNGTKLLQDLGSRSEDEVTEQFGVNHGIPKGEDQTLGKSD 824
QY 751 -----KTKHD-----IYDIDTFDTF-----GDGTPISINA 776
D  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 825 AIPNIGPFGIGISTEESHEEGHKNQALSTVDEPELSDTQLHEDTKENDKLPLESST 884
QY 777 NINEQSG-KOTSN-----GN-----SETSDSPVSHPESDA--AINVEKL 815
D  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 885 ITSPTEGSDTEETPSISEAPKNEQKRDSDSLSKISVPENSRPETAKTSLNLKL 944
QY 816 SGD----- 820
D  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 945 KGDVDISMPKAVIGSSPNDNINVTQDNISGVNSKPLSDDVDPDKNHEEVEKHTNSDN 1004
QY 821 SSETRGILDIN-----DPSVTNNVEH-----DASNTQGSVNTSDI 858
D  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1005 VQSGGGIVNNVKEKELKDTLENPSSSLDGBKAHEVLSEPNLSSDQDMSNTPGPDNTSEE 1064
QY 859 TN--GHSSESLNR-----TTNAQDIKIGRSNEQSDN---QENSSHSSDNGSLTI 904
D  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1065 TTERISNNEKVNREBERLTUKYEDIVLKSHNRESDDGELYDENSLSSTVDESEDA 1124
QY 905 -GQVPSDNTQNTY-----DSQNPHRDTPNALA-----SLP-----SDDKINE-- 941
D  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1125 EAKMGNDTSEMHSNQHIESDQKNDMTKVGDLGTHVQNEISVPVTGEIDELRESK 1184
QY 942 -----IEGFSRSENGRGDTTSTNTHVR-----RTNIVSERRVNSH 979
D  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1185 ESKIHAEERLSHTDIHKINPEDRNSNTLHLKDIRNEENERHLTNQINISQERDLQKH 1244
QY 980 DFIRNGMANNNAHQYITO-----IENNGIIRGOEBSAGHSVNYKDN-----PKESNF 1027
D  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1245 GF-----HTMANLHGDSVRSQINSHSHGQRDRGSGSNVNLNRSNNFNPIPSYNL 1300
QY 1028 SSENDHKNIQVNSRDTKRVREIILKSKONKNNEYSMEYCTYSDERNSSPGPCSRREE 1087
D  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1301 ---YDKKLDLDLYENRNDSTTKELIKKLAIEKNCENBISVKYCDHMTHEBIPLKCTKEK 1357
QY 1088 RKKLCCOISDYCLKYNFYSTIEYVNCIKSEIKSPYKCFKSEGS 1132
D  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1358 TRNLCCAVSDYCMSYFTYDSEYNYCTKREFDDPSYTCFRKAPFS 1402

RESULT 28
Q9N9G9_PLARE
ID Q9N9G9_PLARE PRELIMINARY; PRT; 1433 AA.
AC Q9N9G9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Erythrocyte binding protein (Fragment).
GN Name=EBP;
OS Plasmodium reichenowi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=5854;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Congolesse;
RX MEDLINE=21356926; PubMed=11463469; DOI=10.1016/S0166-6851(01)00298-5;
RA Owara H., Kocken C.H.M., Thomas A.W.;
RT "Molecular characterization of erythrocyte binding protein of
RT Chimpanzee malaria parasite Plasmodium reichenowi.";
RL Mol. Biochem. Parasitol. 116:81-84 (2001).
```

```
DR EMBL; AJ251848; CAB96159.1; -, Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy binding.
DR InterPro; IPR008182; Duffy binding.
DR Pfam; PF05424; Duffy binding; 2.
DR PROSITE; PS00387; PBASE; UNKNOWN 1.
FT CHAIN 9 >1433 erythrocyte binding protein.
FT NON_TER 1 1
FT NON_TER 1433 1433
FT SEQUENCE 1433 AA; 166648 MW; 733D0CDB6EDC27AE CRC64;

Query Match 23.4%; Score 1514.5; DB 2; Length 1433;
Best Local Similarity 28.6%; Pred. No. 1.9e-55;
Matches 406; Conservative 206; Mismatches 451; Indels 357; Gaps 41;

QY 45 NSPELNHHK-TNIYDSYEDVN-----NKLINSFVENKSVKKKSLSPINN--- 90
D  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12 NEYDIEKENKFLVYKKEFNELDKKYGNVQKTDKKIFSFENK-----LDILNNSKS 64
QY 91 --KTKSY--DIIPPSYSYRNDKFNLSLENEDNSGNTSNFNANTSEISIGKDNQ--- 141
D  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 65 NKGWKSYGTPDNIDKMSLIN-KYNN-----KETFNNNYQSFLSTS--SLIKNRYVPIIN 116
QY 142 -----YTFIQKRTH-----LFCAGIKRKSIKWIC-RENSEKITVCVPDRKIQ 182
D  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 117 AVRVSRIILSFLDSKINNGRNTSSNDGLRNCREKRGKNEWDCCKKNGTSNVYVCIQDRIQ 176
QY 183 LCVANFLNRLTMEKFEKIFLISVNTAEKLLNKNKDGKPSIFCNELRNSPDRSSFI 242
D  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 177 LCTVNLISIIKTYTKETMKDHFIEASRRSQLLKKNNDNNYSKFCNDLKNFLDYGHAM 236
QY 243 GDDMDFGGNTDRVKGVIINTKFSDYKKNVEKLNKIKKEWWEKKNANLWHMIVNKHGNI 302
D  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 237 GNDMDFGGSTKAEKNIQEVFKGVHGKISEHEIKFNFRKKNWNEFREKLWEMAMISEKNNL 296
QY 303 SKECAIIPAEPEQINLIWKENENFLMEKKRFLNFKDCVENKRYEACFGGRLPCSSY 362
D  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 297 S-DCKNIPEELQINQIKWEHGEFLSERNNRLKPKSKCNNTLYEACAKECIDPCMKY 355
QY 363 TSPMKSKTQMEVLTNLYKKKN-SGVYDKNFLNDLFKNKNKNDLDDFF-KNEKEYDDLCD 420
D  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 356 RDWIIKSKFEWHTLSKEYERKNVSNKDAEKLK-FSKNDAKVSLLLDKDAEYSKYCD 414
QY 421 CRYTATIIKSFNLGPAKNDVDIASQINVNDLRLGFGCNYKS--NNEKSWNCTGTFTNKPFG 478
D  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 415 CKHTTLVKSVLNGKONTKEERETIDDDFSKFGCDKNSVDNTRKEWECKPYILSTKD 474
QY 479 TCPEPRRQTLCLGRTYLLHRGHEEDYKEHLGLGASIIYEAQLLKYYKEKDNALCSIIQNS 538
D  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 475 VCVPPRRQELCLGNIDRIYDKNLLMIKEHILATAIYESRLKRYKKNKDDNEVCNVINKS 534
QY 539 YADLADIIKGGDIIKDYKGMENLNKVNKKRNEESLKI PREKWNDEKENVKNVMS 598
D  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 535 FADIRDIIKGGDYWNLNRLKVGKINTNSKYVHKKNENDKLFRDEWNNVKKDWNVMS 594
QY 599 AVLKNKETCKDYDFOKIPOFLRWFKWGDGDFCEKKEKIYSPESPKVECKKDCDENTC 658
D  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 595 WPFKDKTVCKE-DDIENIQPFWFSEWGGDDYQ---YKLMIDTLKACEEKGYYDITC 650
QY 659 KKKSEYKQWIDLKSEYKQVDKYT--KDKNKKMYNDIVKKNKEANVYLKSKSECKXD 716
D  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 651 KKKCSYKDWISKQKLYDKQVTIYKEYQRNNRIY---PIVKTMEPKVYLKEYSKKCSN 707
QY 717 VNEFDKIFNESPNEYEDMC-----KKCDSEIKVINEIKYP 750
D  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 708 INFEDFNEBSEVSHSDYKKNKCMISTKVLDPVIPAENKNTVEALETVVPKENTIERKNESDTE 767
QY 751 KTKH-----DIYDIDTFSDTFDGTGTPIS-----INANI---NEQSQS 784
D  : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 768 EQKNQEIPINNYLKDSQQRVGENGCTKYSSQBEDLGRSNEEGVTOEFLGNDIIPKGEQISG 827
```



```

QY 785 KD-----TSNTGSETSDSPVSHE-----PES 806
Db 828 KSGVIONRSEMGISTSESRHEEGHNEQTLSVDQSELSDTLQSHEDTKENDFLA 887
QY 807 DAAINVEKLGDESSETRGILDNDPSVTNNVEHDASNTQGSVNTSDITNGHSESS 866
Db 888 SSTTSPMESGSSDTEETQ--SISEGPKGNEQKRDSDSSKISVSPEISRTETHDKOT 944
QY 867 LNRITTAQDIK-----IGRSGNEQSDNQENSSH--SSDNGSGSLTIGQVPS- 910
Db 945 SNLLRLKEDVDISMPKAVIGKDPNDHINNTAEGDHISGVNSPLSDGVRPKNHEEVKEH 1004
QY 911 -----DNTQNTVDS-----QNPRTD- PNALASL-----PS 935
Db 1005 TSNPDNVQSGENVMNVENQDKTLENTSSSLGEGKAHELSBPNLSSGQDMSNTAPL 1064
QY 936 DD-----KINE-----LEGFDSRDSENG-----954
Db 1065 DNTSEETTERISNNEYKVNREDERTLTKYEDIVLKSHMNRSDDDLDGKNSDVPTVN 1124
QY 955 -----RGDTTSTNTHOV-----965
Db 1125 DESEDEVEKMRDNTSEISESSQHSIQDQENDKTVDGLGTHVQNEINDSVTGENKE 1184
QY 966 -----RRTNIVSRRVNSHDFIRNGMANNAH-----H 993
Db 1185 QTVHDEGVRLSHKDKKHINLEDKNSLILHLKMRNEERKQLANQINISQQRDLQEH 1244
QY 994 QYIT--QIENNGIITRQGESAGNSVNYKONPKRS-----NFSSEN-----DHK 1034
Db 1245 GFHTWNNLHGVSQSGSINHGNPKRQDRGKPSGNVLNMGSNNNNFHIFSRNLYDKK 1304
QY 1035 KNIQYNSRDKRVREKIIKLSKONKNNEYSMEYCTYSDRNSSPGCSREERKLCQ 1094
Db 1305 LNLQLYNRNDSTYKELIKKLAINKCEISAKYCSMDIKDEPLKTCCTEKTINLCCA 1364
QY 1095 ISDYCLKYNFYNSIYNYNCIKSEIKSPYKCFKSEGOSSI 1134
Db 1365 VADYCLSYFTYDSKEYSNCTREFEDPSYTCFRKKGFSSM 1404

```

RESULT 29

```

QY Q81796 PLAFPA PRELIMINARY; PRT; 614 AA.
Db AC Q81796;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Erythrocyte binding antigen region II (Fragment).
GN Name=eba-175;
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22588500; PubMed=12702678;
RA Baum J., Thomas A.W., Conway D.J.;
RT "Evidence for diversifying selection on erythrocyte-binding antigens
of Plasmodium falciparum and P. vivax.";
RL Genetics 163:1327-1336(2003).
DR EMBL; AJ438810; CAD27535.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy binding.
DR Pfam; PF05424; Duffy binding; 1.
FT NON_TER 614 614
SQ SEQUENCE 614 AA; 73292 MW; A168F604AB322F7F CRC64;

```

Query Match 18.7%; Score 1210.5; DB 2; Length 614;
 Best Local Similarity 38.2%; Pred. No. 4.3e-43;
 Matches 245; Conservative 112; Mismatches 239; Indels 45; Gaps 11;

```

QY 151 LPACGIKRSIKWICRESEKIT--VCVPDRKIQLCVANFLNSRLTMEKPEIFLISVNT 209
Db 11 LSSCREKRGKMKDCKKNDNSYVCIPDRIQICIVNLSIIKTYTETKMDHFEASKK 70
QY 210 EAKLLYNNEGKOPSIFCNELNRNFSDFRSSFIDGMDFGNTDRVKGYINTKSDYYKE 269
Db 71 ESQALLKNDNKNYSKFCNDLKNFLDYGHLAGNDMDFGGYSTKAENKIQEVFKGAHGK 130
QY 270 KNVSEKLANNIKKEWKEKNNANLWNNHKNISKECAIIPAEPPQINLWIKENNENFLM 329
Db 131 ISEHEIKNFKRKNWNEPREKLEWALWSEHKNNIN--NKNIPQELQITQWIKWHGEFLL 189
QY 330 EKKEFLNLNIDKQCVENKYEACFCGCRLLPCSSYTSFMKSKSTQMEVLNLYKKNSGVYDK 389
Db 190 ERDNRSKLPKSKCKNNNTLYEACEKECIDPCMKYRDWIIRSKFEWHTLSKEYETQN--VSK 247
QY 390 NNFNDLFPKNNKNDLDDFFKN--EKEYDDLDCCRYTATIIKSFNGPAKNDVDIASQIN 447
Db 248 ENAENYLIKKMMDAKVSLNLCDAEYSKYCDCKHTTTLVKSVLNGNDNTIKEKREHID 307
QY 448 VNDLRGFGCNYS--NNEKSNCTGTPTNKPFGTCEPPROTLCIGRTYLLHRGHEEDYK 505
Db 308 LDDPSKFGCDKNSTVDTTKVWECKPKYKLTQKVCVPPRQELCLGNIDRIYDKNLLMIK 367
QY 506 EHLIGASIYEAQLKYYKEKDNALCSIIQNSYADLADIIGSDIISKDYKCKMEENLN 565
Db 368 EHILAIAIYSRILKRYKKNCKDEKCKIINKTFADIRDIIGGTDYNDLSNRKLVGKIN 427
QY 566 KVNCKKGRNEESLIPREKWDENKENVWKSVAVLKNKCTCKDYDFQKIPQFLRWPK 625
Db 428 TNSYVHRNKQNDKLPDANWVKIKDQVNVISWVFKDKTVCKE--DDIENIPQFFRWFSE 486
QY 626 WGDPPCEKREKIKYSPSPFVECKKDCDENTCKNCKSEYKKWIDLKSEYEKQV---DK 682
Db 487 WGDYCYQDKTKMI---ETLKVCECKEPCEDDNCCKNSKYKESWISKKEETNKAQKQ 543
QY 683 YTKDNKKMVDNIDVENKKEANVYLKESKECKDQNFDDKIFNESPNEBYEDMCKCKDBIK 742
Db 544 YQKGNVYKMY---SEPKSIKPEVYLKYSKCKSNLNFEDPEKEELHSDYKKNKCTMCPVK 600
QY 743 YLNEIKYPTKTHDIYDITFSDTFTGDTPTPISANINANEQOS 783
Db 601 -----DVPISIRN-NEQTS 614

```

RESULT 30

```

QY Q81791 PLAFPA PRELIMINARY; PRT; 616 AA.
Db AC Q81791;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Erythrocyte binding antigen region II (Fragment).
GN Name=eba-175;
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22588500; PubMed=12702678;
RA Baum J., Thomas A.W., Conway D.J.;
RT "Evidence for diversifying selection on erythrocyte-binding antigens
of Plasmodium falciparum and P. vivax.";
RL Genetics 163:1327-1336(2003).
DR EMBL; AJ438822; CAD27547.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy binding.
DR Pfam; PF05424; Duffy binding; 1.
FT NON_TER 616 616
SQ SEQUENCE 616 AA; 73292 MW; A168F604AB322F7F CRC64;

```

```
SQ SEQUENCE 616 AA; 73559 MW; 4B3BE72152B7D1F0 CRC64;

Query Match
Best Local Similarity 18.7%; Score 1209.5; DB 2; Length 616;
Matches 253; Conservative 115; Mismatches 236; Indels 69; Gaps 15;

QY 121 NTNSNNPANTSEISIGKDNKQYTFIQKRTHLFAGGIRKRSIKWICRENSKIT-VCPDPR 179
Db 3 NTSSNN-----EV-----LSSCREKRGKMKWDCKKNDNSNYVICPDR 40

QY 180 KIQLCVANFLNSRLTMEKFEKIFLISVNTAEKLLYNKNEGKPSIFCNELRNSFSDFRS 239
Db 41 RIQLCIVNLSIIKTYTKETMKDHFIEASKESQLLLKNDNEYSKFCNDLKNSFLDYGH 100

QY 240 SFTGDDMDPGNTDRVKGINTKPSDYKKEKNNKANNIKKEWKEKNAKLNHMHVNHK 299
Db 101 LAMGNDMDFGYSTKAENKIQEVFGAHSIEHKKNFREKWNBEPREKLWEAMLSEHK 160

QY 300 GNISKECALIIPABEPOINLWIKENENFLMEKRLFLNIDKCKVENKKYBACGGCRLPC 359
Db 161 NNIN-NCKNIPQBELQITQWIKEMHGFLLERNRSKLPKSKCKNNTLYACEKECIDPC 219

QY 360 SSVTSFMKSKTOMEVLNLYKKNSGVNDKNNFLNDFK-KNNKND--LDDFFKN-EKEY 415
Db 220 MKYRDWIIRSKFEWHLSKEYETQK--VPKNAENYLIKISEKNQDAKVSLLLNNDCAEY 277

QY 416 DDLCDCRYTATIKSFLNGPAKNDVDIASQINVDNLRGFCNYS--NNEKSNCTGTFT 473
Db 278 SKYDCCKHHTTLVKSVLNGNDNTIKEREHIDLDDFSFGCDKNSVDNTNTKWCKNPIYI 337

QY 474 NKPPGTCPPRRQTLCLGRITLHGRHEEDYKEHLGASIEAQLLYKYKEKDENA LCS 533
Db 338 LSTKDVCPVPRQELCIGNIDRIYDKNLLMIKEHILAIAIESRLKRYKNKDDKEVCK 457

QY 594 WKYMSAVLNKNETCKDYDFQKIPQFLRWPKFGWDDFCERKEKIKYSFESFKVECKKDC 653
Db 458 WNVISWFKDKTVCKE-DDIENIPQFRWPFSEWGDDYCDQKTKMI---ETLKVECKEKP 513

QY 654 DENTCKNCKSEYKWKIDLKSEYEQV---DKYTKDNKKNMYDNIDEVKNEANVYLKEK 710
Db 514 EDDNCKSKNSYKEWISKKEEYNKQAKQYQYQKGNKYKMY---SEFKSIKPEVILKY 570

QY 711 SKECKDVNFDDKIPNESPNEDMCKKCDKIKYLNEIKYKPKTKHDIYDIDTFSDFGDT 770
Db 571 SEKCSNLNFEDEFKEELHSDYKKNCTWCPEVK-----DV 604

QY 771 PISINANINEQQS 783
Db 605 PISIIRN-NEQTS 616

RESULT 31
Q81797_PLAFA PRELIMINARY; PRT; 616 AA.
AC Q81797;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Erythrocyte binding antigen region II (Fragment).
GN Name=eba-175;

OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22588500; PubMed=12702678;
RA Baum J., Thomas A.W., Conway D.J.;
RT "Evidence for diversifying selection on erythrocyte-binding antigens
of Plasmodium falciparum and P. vivax.";
```

```
RL Genetics 163:1327-1336(2003).
DR EMBL; AJ438808; CAD27533.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duff binding.
DR Pfam; PF05424; Duff binding; 1.
FT NON_TER 1
FT NON_TER 616
SQ SEQUENCE 616 AA; 73589 MW; 312B7CF98D9796BE CRC64;

Query Match
Best Local Similarity 18.7%; Score 1209.5; DB 2; Length 616;
Matches 254; Conservative 112; Mismatches 238; Indels 69; Gaps 15;

QY 121 NTNSNNPANTSEISIGKDNKQYTFIQKRTHLFAGGIRKRSIKWICRENSKIT-VCPDPR 179
Db 3 NTSSNN-----EV-----LSSCREKRGKMKWDCKKNDNSNYVICPDR 40

QY 180 KIQLCVANFLNSRLTMEKFEKIFLISVNTAEKLLYNKNEGKPSIFCNELRNSFSDFRS 239
Db 41 RIQLCIVNLSIIKTYTKETMKDHFIEASKESQLLLKNDKYNKSKFCNDLKNSFLDYGH 100

QY 240 SFTGDDMDPGNTDRVKGINTKPSDYKKEKNNKANNIKKEWKEKNAKLNHMHVNHK 299
Db 101 LAMGNDMDFGYSTKAENKIQEVFGAHSIEHKKNFREKWNBEPREKLWEAMLSEHK 160

QY 300 GNISKECALIIPABEPOINLWIKENENFLMEKRLFLNIDKCKVENKKYBACGGCRLPC 359
Db 161 NNIN-NCKNIPQBELQITQWIKEMHGFLLERNRSKLPKSKCKNNTLYACEKECIDPC 219

QY 360 SSVTSFMKSKTOMEVLNLYKKNSGVNDKNNFLNDFK-KNNKND--LDDFFKN-EKEY 415
Db 220 MKYRDWIIRSKFEWHLSKEYETQK--VSKENAENYLIKISEKNDAKVSLLLNNDCAEY 277

QY 416 DDLCDCRYTATIKSFLNGPAKNDVDIASQINVDNLRGFCNYS--NNEKSNCTGTFT 473
Db 278 SKYDCCKHHTTLVKSVLNGNDNTIKEREHIDLDDFSFGCDKNSVDNTNTKWCKKPYK 337

QY 474 NKPPGTCPPRRQTLCLGRITLHGRHEEDYKEHLGASIEAQLLYKYKEKDENA LCS 533
Db 338 LSTKDVCPVPRQELCIGNIDRIYDKNLLMIKEHILAIAIESRLKRYKNKDDKEVCK 397

QY 534 IIONSVADLADIIGSDIIDYKYGKMEENLNKKNKKNESLKFREKKNWDEKENV 593
Db 398 IINKTFADIRDIIGTDYNDLSNRKLGVKINTNSYVHRNKQNDKLFREDWVKVIKDV 457

QY 594 WKYMSAVLNKNETCKDYDFQKIPQFLRWPKFGWDDFCERKEKIKYSFESFKVECKKDC 653
Db 458 WNVISWFKDKTVCKE-DDIENIPQFRWPFSEWGDDYCDQKTKMI---ETLKVECKEKP 513

QY 654 DENTCKNCKSEYKWKIDLKSEYEQV---DKYTKDNKKNMYDNIDEVKNEANVYLKEK 710
Db 514 EDDNCKSKNSYKEWISKKEEYNKQAKQYQYQKGNKYKMY---SEFKSIKPEVILKY 570

QY 711 SKECKDVNFDDKIPNESPNEDMCKKCDKIKYLNEIKYKPKTKHDIYDIDTFSDFGDT 770
Db 571 SKKCSNLNFEDEFKEELHSDYKKNCTWCPEVK-----DV 604

QY 771 PISINANINEQQS 783
Db 605 PISIIRN-NEQTS 616

RESULT 32
Q817A0_PLAFA PRELIMINARY; PRT; 614 AA.
AC Q817A0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Erythrocyte binding antigen region II (Fragment).
GN Name=eba-175;
```



```

Db 514 EDDNCKSKNSYKEMWISKKKEYNKQAKQYQYQKGNMYKMY---SEFKSIKPEVYLKXY 570
Qy 711 SKECKDVNFDKIFNESPNEYEDMKCKDEIKYLNELKYPKTKHDIYDIDTFSDTFDGT 770
Db 571 SEKSNLNFDEFEKELHSDYKQKCTMCPVK-----DV 604
Qy 771 PISINANEQOS 783
Db 605 PISIIRN-NEQTS 616

RESULT 34
Q25739_PLAFA
ID Q25739_PLAFA PRELIMINARY; PRT; 616 AA.
AC Q25739;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Erythrocyte binding antigen region II (Fragment).
GN Name=EBA-175;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=VI-9;
RX MEDLINE=97237565; PubMed=9084043; DOI=10.1016/S0166-6851(96)02791-0;
RA Liang H., Sim B.K.;
RT "Conservation of structure and function of the erythrocyte-binding
domain of Plasmodium falciparum EBA-175."
RL Mol. Biochem. Parasitol. 84:241-245(1997).
DR EMBL; U27396; AAB51620.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy binding.
DR Pfam; PF05424; Duffy binding; 1.
FT NON_TER 1 616
FT SEQUENCE 616 AA; 73517 MW; 2C1B4990CF732693 CRC64;

Query Match 18.6%; Score 1206.5; DB 2; Length 616;
Best Local Similarity 37.7%; Pred. No. 6.3e-43;
Matches 255; Conservative 115; Mismatches 231; Indels 75; Gaps 17;

Qy 121 NTNSNFPANTSBEISIGKDNKQYTFIQKRTHLFACGIRKSIKWTCRENSEKIT-VCVPDR 179
Db 3 NTSSNN-----EV-----LSNCREKRGKMKWCKKKNDRSNYVCIPDR 40

Qy 180 KIQLCVANFLNSRLTWEKPEIFLISVNTVEAKLLYNKNEGKDPSIFCNELRNSFSDPRS 239
Db 41 RIQLCVNLNLSIIKTYTETWKDHFIEASKESQLLLKNDKNKNSKFCNDLKNFLDYGH 100

Qy 240 SFIGDDDFGNTDRVGYINTKFSYVYKEQVKEQVKEQVKEQVKEQVKEQVKEQVKEQV 299
Db 101 LAMGNDMDFGYSYKAENKQYEVFGAGHGEISEHEIKNFRKWNNEFKLWEAMUSEHK 160

Qy 300 GNISKECAIIPAEPPQINLMIKWENFLMEKRLFLNIKDCKVENKYEACGGCRLPC 359
Db 161 NNIN-NCKNIPQELQITQMIKEWHGEFLERDNRSLPKSKCKNNTLYEACEKCIDPC 219

Qy 360 SSYTSFMKSKTQMEVLTNLYKKNKSGVDKNNFLNDLPK-KNNQD--LDDFFQN-EKEY 415
Db 220 MKYRDWTIRSKPWHHTLSKEYETQN--VSKENAEYLIKISKNDKDKVSLLLNCDAY 277

Qy 416 DDLCDCRYATIIKSLFNGPAKNDVDIASQ---INVNDLRGFCGNYS--NNEKSNWCTG 470
Db 278 SKYCDCKHTTTLVKSVLNG---NNTYKEEREHIDDDFSKFGCKNSVDNTNWKVECKN 334

Qy 471 TFTNKFPGCTEPPRQTLCLGRYTLHLRGHBEDEYKEHLLGASYYEAQLLYKYKEKEDENA 530
Db 335 PYLSTKDVCPVPRRQELCLGNIDRIYDKNLLMIKEHILALAIYESILRKRYKNKDDKE 394

```

```

Qy 531 LCSIIQNSVADLADIIGSDIIKDYGYKQVKEQVKEQVKEQVKEQVKEQVKEQVKEQV 590
Db 395 VCKIINKTADIRDIIGTGYNDLSNRKLGVKINTNSYVHRNKKNDKLFDEWKKVIX 454
Qy 591 ENWKVMSAVLKNKCTCKDYDFOKIPQFLRWPFKWGGDDFCCEKREKIKYSFESFKVECKK 650
Db 455 KDVWNVISWVFKDKTVCKE-DDIENIPQFRWFWSEWGGDYCY---DKTKTIELKVECKE 510
Qy 651 KDCDENTCKNCKSEYKWKWIDLKKESEYKQV---DKYTKDNKKNQYDNIDEVKNEANVYL 707
Db 511 KPCEDDNCCKSKNSYKEMWISKKKEYNKQAKQYQYQKGNMYKMY---SEFKSIKPEVYL 567
Qy 708 KEKSKCKDVNFDKIFNESPNEYEDMKCKDEIKYLNELKYPKTKHDIYDIDTFSDTFG 767
Db 568 KKYSEKSNLNFDEFEKELHSDYKQKCTMCPVK----- 602
Qy 768 DGTPIISINANEQOS 783
Db 603 -DVPISIIRN-NEQTS 616

RESULT 35
Q25736_PLAFA
ID Q25736_PLAFA PRELIMINARY; PRT; 614 AA.
AC Q25736;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Erythrocyte binding antigen region II (Fragment).
GN Name=EBA-175;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ENGL;
RX MEDLINE=97237565; PubMed=9084043; DOI=10.1016/S0166-6851(96)02791-0;
RA Liang H., Sim B.K.;
RT "Conservation of structure and function of the erythrocyte-binding
domain of Plasmodium falciparum EBA-175."
RL Mol. Biochem. Parasitol. 84:241-245(1997).
DR EMBL; U27393; AAB51617.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy binding.
DR Pfam; PF05424; Duffy binding; 1.
FT NON_TER 1 614
FT SEQUENCE 614 AA; 73334 MW; B1994894CE1902C3 CRC64;

Query Match 18.6%; Score 1204.5; DB 2; Length 614;
Best Local Similarity 38.2%; Pred. No. 7.7e-43;
Matches 244; Conservative 110; Mismatches 239; Indels 45; Gaps 11;

Qy 154 CGIKRKSIIKWTCRENSEKIT-VCVPDRKTLQCVANFLNSRLTWEKPEIFLISVNTVEAK 212
Db 14 CREKRGKMKWCKKKNDRSNYVCIPDRRIQLCVNLNLSIIKTYTETWKDHFIEASKESQ 73

Qy 213 LLYNKNEGKDPSIFCNELRNSFSDRSFSGDDMDFGNTDRVGYINTKFSYVYKEQV 272
Db 74 LLLKNDKNKNSKFCNDLKNFLDYGHLAGNDMDFGYSYKAENKQYEVFGAGHGEISEHEIK 133

Qy 273 EKLNNIKKWEKWKANLWNHMIIVNHKGNISKECAIIPAEPPQINLMIKWENFLMEKK 332
Db 134 HEIKNFRKWNNEFKLWEAMUSEHKNNIN-NCKNIPQELQITQMIKEWHGEFLERD 192

Qy 333 RFLNLIKDKCVENKYEACGGCRLPCSSYTSFMKSKTQMEVLTNLYKKNKSGVDKNNF 392
Db 193 NRSKLPKSKCKNNTLYEACEKCIDPCMKYRDWTIRSKPWHHTLSKEYETQN--VSKENA 250

Qy 393 LNDLPKKNKNDLDDFFQN--EKEYDDLCDCRYATIIKSLFNGPAKNDVDIASQINVD 450

```

```

Db 251 ENYLKKGQNDKAVSLNLLNCDASYSKYCDCKHTTTLVKSVLNGNDNTIKREKREHIDLD 310
Qy 451 LRGFGCNYS--NNEKSNWCTGTNTKPPGCEPRQTLCLGRYTLHLRHEEDYKHL 508
Db 311 FSKGCDKNSVDYTKVWECKPKYKSTYDVCVPRQELCLGNDRIYDKNLLMIKEHI 370
Qy 509 LGASIYEAQLKYKKEKDNALCSIIIONSADIIKSDIIKDYGGKQKQENLNKN 568
Db 371 LAIAIYBSRIILKRYKNDKDEVCCKIINKTPADIRDIIGGTYWDLNLRKLVGKINTS 430
Qy 569 KDKRNEESLAKFREKWDENKENVKMSAVLNKTKCKDYDKFQKIPQLRWFKEGD 628
Db 431 NYVRHNKENDKLFKRDAMWVKIKDQVNVISWVFKDKTVCKE--DIENIPQFRWFSEGD 489
Qy 629 DFCRKRKEKYSFESFKVECKKDCDENTCNKCSYKWKWIDLKKSSEYKQV---DKYTK 685
Db 490 DYCDQKTKMI---ETLKVCEKEKCEDNCKSKNSYKESKKEEYNKQAKQYQVOK 546
Qy 686 DNKNKMYNDIDDEVKNKEANVYLKSKCEKDVNFDDKI FNSPNEYEDMCKCKDEIKYLN 745
Db 547 GNNYKMY---SEFKSIKPEVLYKYSKCSNLNFEDEBPKBELHSDYKKNKCTMCPVK--- 600
Qy 746 EIKPYKTKHDIYDITFSDTGDGTPISINANINEQOS 783
Db 601 -----DVPISIRN-NEQTS 614

RESULT 36
Q8I790_PLAFA PRELIMINARY; PRT; 614 AA.
AC Q8I790;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Erythrocyte binding antigen region II (Fragment).
GN Name=eba-175;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22588500; PubMed=12702678;
RA Baum J., Thomas A.W., Conway D.J.;
RT "Evidence for diversifying selection on erythrocyte-binding antigens
of Plasmodium falciparum and P. vivax.";
RL Genetics 163.1327-1336(2003).
DR EMBL; AJ38824; CAD27549.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0009405; P:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy binding.
DR Pfam; PF05424; Duffy_binding; 1.
FT NON TER 614
FT NON TER 614
SQ SEQUENCE 614 AA; 73306 MW; E7E6CD3C48774839 CRC64;

Query Match 18.6%; Score 1204.5; DB 2; Length 614;
Best Local Similarity 38.1%; Pred. No. 7.7e-43;
Matches 244; Conservative 112; Mismatches 240; Indels 45; Gaps 11;

Qy 151 LPACGIRKRSIKWICRENSKIT--VCVPRDRIQLCVANFLNRLTWEKFEIPLISVNT 209
Db 11 LSSCREKRGKMKWCKKNDKNSVCIPIRDIQLCVNLISIIKYTKETMKEDHIEASKK 70
Qy 210 EAKLLYNNEKDPISIFCNELRNFSDFRSSFIDGMDPGNTRVKGVIKTFSDYKKE 269
Db 71 ESQLLKKNKNSKFCNDLKNFLDYGLHAGNDMDFGYSTKAENKIQEVFKGAGK 130
Qy 270 KQVEKLANIKKEWKEKKNAMLNHMI VNHKNISKECAIIPAEPOINLWIKENENFLM 329
Db 131 ISEHEIKNFRKWNNEFKELWEAMLSEHKNNIN--NCKNIIPQBELQITQWIKWHGEFL 189
Qy 330 EKRLFLNIDKQVKNKKYKACFGCRLPCSSYTSFMKSKSTQMEVLTLNLYKKQNSGVDK 389

```

```

Db 190 ERDNRSLPKSKCKNNTLYEACEKECIPCKMYRDMWIRSKFEWHITLSKEYETQN--VSK 247
Qy 390 NNFLNDLFKQNNKNDLDDFFQN--EKSYDDLCDCRYTATIIKSPFNGPAKNDVDIASQIN 447
Db 248 ENAENYLIKKQNDKAVSLNLLNCDASYSKYCDCKHTTTLVKSVLNGNDNTIKREKREHID 307
Qy 448 VNDLRGFGCNYS--NNEKSNWCTGTNTKPPGCEPRQTLCLGRYTLHLRHEEDYK 505
Db 308 LDDFSKFGCDKNSVDYTKVWECKPKYKSTYDVCVPRQELCLGNDRIYDKNLLMIK 367
Qy 506 EHLGASIYEAQLKYKKEKDNALCSIIIONSADIIKSDIIKDYGGKQKQENLNKN 565
Db 368 EHILAIYBSRIILKRYKNDKDEVCCKIINKTPADIRDIIGGTYWDLNLRKLVGKIN 427
Qy 566 KVNKDKRNEESLAKFREKWDENKENVKMSAVLNKTKCKDYDKFQKIPQLRWFKE 625
Db 428 TNSNYVRHNKENDKLFKRDAMWVKIKDQVNVISWVFKDKTVCKE--DIENIPQFRWFSE 486
Qy 626 WGDGDFCERKRKEKYSFESFKVECKKDCDENTCNKCSYKWKWIDLKKSSEYKQV---DK 682
Db 487 WGDYCDQKTKMI---ETLKVCEKEKCEDNCKSKNSYKESKKEEYNKQAKQYQVOK 543
Qy 683 YTKDNKMYNDIDDEVKNKEANVYLKSKCEKDVNFDDKI FNSPNEYEDMCKCKDEIKYLN 742
Db 544 YQGNMYKMY---SEFKSIKPEVLYKYSKCSNLNFEDEBPKBELHSDYKKNKCTMCPVK 600
Qy 743 YLNEIKPYKTKHDIYDITFSDTGDGTPISINANINEQOS 783
Db 601 -----DVPISIRN-NEQTS 614

RESULT 37
Q8I0C8_PLAFA PRELIMINARY; PRT; 616 AA.
AC Q8I0C8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Erythrocyte binding antigen region II (Fragment).
GN Name=eba-175;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22588500; PubMed=12702678;
RA Baum J., Thomas A.W., Conway D.J.;
RT "Evidence for diversifying selection on erythrocyte-binding antigens
of Plasmodium falciparum and P. vivax.";
RL Genetics 163.1327-1336(2003).
DR EMBL; AJ38814; CAD27539.1; -; Genomic DNA.
DR EMBL; AJ38820; CAD27545.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0009405; P:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy binding.
DR Pfam; PF05424; Duffy_binding; 1.
FT NON TER 616
FT NON TER 616
SQ SEQUENCE 616 AA; 73600 MW; A0347CAFA5214D47 CRC64;

Query Match 18.6%; Score 1204.5; DB 2; Length 616;
Best Local Similarity 37.6%; Pred. No. 7.7e-43;
Matches 253; Conservative 113; Mismatches 238; Indels 69; Gaps 15;

Qy 121 NTNSNPNANTSEISIGKNDKQYTFIQKRTHLFACGIRKRSIKWICRENSKIT--VCVPRD 179
Db 3 NTSNN-----EV-----LSNCREKRGKMKWCKKNDKNSVCIPIR 40
Qy 180 KIQLCVANFLNRLTWEKFEIPLISVNTAEKLLYNNEKQDPSIFCNELRNFSDFRS 239
Db 41 RIQLCVNLISIIKYTKETMKDHPFEASKESQLLKNKNDKNSKFCNDLKNFLDYGH 100

```



```

RESULT 39
Q27248 PLAF A PRELIMINARY; PRT; 616 AA.
AC Q27248;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Erythrocyte binding antigen region II (Fragment).
GN Name=EBA-175;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
NUCLEOTIDE SEQUENCE.
RC STRAIN=FVO, 106/10, and FCB;
RX MEDLINE=97237565; PubMed=9084043; DOI=10.1016/S0166-6851(96)02791-0;
RA Liang H., Sim B.K.;
RT "Conservation of structure and function of the erythrocyte-binding
domain of Plasmodium falciparum EBA-175."
RL Mol. Biochem. Parasitol. 84:241-245(1997).
DR EMBL; U27390; AAB51614.1; -; Genomic DNA.
DR EMBL; U27385; AAB51610.1; -; Genomic DNA.
DR EMBL; U27389; AAB51613.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy binding.
DR Pfam; PF05424; Duffy binding; 1.
FT NON_TER 1 616
FT NON_TER 616
SQ SEQUENCE 616 AA; 73669 MW; A19ECD04BF908F7F CRC64;

Query Match 18.6%; Score 1203.5; DB 2; Length 616;
Best Local Similarity 37.6%; Pred. No. 8.5e-43;
Matches 253; Conservative 112; Mismatches 239; Indels 69; Gaps 15;

Qy 121 NTSSNNFANTSEISIGKDNKQYTFIQKRTHLFACGIRKSIKWICRENSEKIT-VCVPDR 179
Db 3 NTSSNN-----EV-----LSNCRKRGKMGWCKKKNDNRNVCIPDR 40

Qy 180 KIQLCVANFLNSRLTMEKFEIPLISVNTAKLLYNKNEKDPISFCNELRNSFSDPRS 239
Db 41 RIQLCIVNLNLSIITTYTETKMDHFIKSKESQLLLKNDKNKYNKSKFCNDLKNFLDYGH 100

Qy 240 SFIGDDMFGNTDRVKGINTKFSDDYKKNVKNLNIKEWKNKKNANLWNIHVK 299
Db 101 LAMGNDMFGGYSTKAENKIQEVFGAGHSEIHEIKNFRKKNWNEFREKLEWMLSEHK 160

Qy 300 GNISKECAIIPABEPQINLMIKEWNEFLMEKRLFLNIKDKCVENKKEACFGGCRIPC 359
Db 161 NNIN-NCKNIPQEBELQITQIKWGHGFLLEDRNRSKLPKSKCKNNTLYEACEKIDPC 219

Qy 360 SSYTSFMKSKTQMEVLTNLYKKNSGVQDNFLNLFK-KNNKD--LDDFFKN-EKEY 415
Db 220 MKYRDWIRSFHWHTLSKEYETQK--VPKENAENYLIKISENKDANKVSLNLCDAEY 277

Qy 416 DDLCDCRTATIKSFLNGPAKNDVDIASQINVDLGRGFCNYKS--NNEKSWNCTGTFT 473
Db 278 SKYCDCHHTTLVKSVLNGNDNTIKKREHLDLDDFSKGCDCNKSVDNTNTKVECKPKY 337

Qy 474 NKFGTCEPPRRQTLCLGRYTLHHRGHEEDYKEHLLGASIEAQLLYKYKBDENALCS 533
Db 338 LSTKDVCPVRQELCLGNDIRIDYKLLMIKEHLAIAIYESILRKRYKNKDDKEVCK 397

Qy 534 IIONSADLADIKGSDIIKDYGGKMEENLNKYNKDKRNEESLKIPREKWDENKENV 593
Db 398 IINKTFADIRDIIGGTDYNDLSNRKLVGKINTNSYVHRNKQNDKLPDEWNVKIKDV 457

Qy 594 WKMSAVLNKTKCKDYKFKQIPQLRWPKEWGGDCEKREKEKIYFSEFKECKKDC 653
Db 458 WNVISWFKQTKVCK-DDIENIPQFRFWSEWGGDYCDQTKWMI---ETLKVCKECPK 513

Qy 654 DENTCKNCSYKKWIDLKSEYEQV---DKYTKDNKKQMDYNDIVENKNEANVYLKEK 710

```

```

Db 514 EDDNCKKCKSYKEWISKKEEYNKQAKQYQYQKGNMYK---SEPKSIKPEVYLKCY 570
Qy 711 SKCKQVNFDDKIPNESPNVEDMKCKDEIKYLNEIKYPKTKHDIDIDITFSDTFGDT 770
Db 571 SEKSNLNFDEFKBEHLSDYKKNKCTWCPEVK-----DV 604
Qy 771 PISINANINEQOS 783
Db 605 PISIIRN-NEQTS 616

RESULT 40
Q27361 PLAF A PRELIMINARY; PRT; 616 AA.
AC Q27361;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Erythrocyte binding antigen region II (Fragment).
GN Name=EBA-175;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
NUCLEOTIDE SEQUENCE.
RC STRAIN=DIV30, and 7G8;
RX MEDLINE=97237565; PubMed=9084043; DOI=10.1016/S0166-6851(96)02791-0;
RA Liang H., Sim B.K.;
RT "Conservation of structure and function of the erythrocyte-binding
domain of Plasmodium falciparum EBA-175."
RL Mol. Biochem. Parasitol. 84:241-245(1997).
DR EMBL; U27388; AAB51612.1; -; Genomic DNA.
DR EMBL; U27386; AAB51611.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy binding.
DR Pfam; PF05424; Duffy binding; 1.
FT NON_TER 1 616
FT NON_TER 616
SQ SEQUENCE 616 AA; 73547 MW; 1B1A5A0A64F1C18B CRC64;

Query Match 18.6%; Score 1202.5; DB 2; Length 616;
Best Local Similarity 37.3%; Pred. No. 9.3e-43;
Matches 253; Conservative 117; Mismatches 236; Indels 69; Gaps 15;

Qy 121 NTSSNNFANTSEISIGKDNKQYTFIQKRTHLFACGIRKSIKWICRENSEKIT-VCVPDR 179
Db 3 NTSSNN-----EV-----LSSCRKRGKMGWCKKKNDNRNVCIPDR 40

Qy 180 KIQLCVANFLNSRLTMEKFEIPLISVNTAKLLYNKNEKDPISFCNELRNSFSDPRS 239
Db 41 RIQLCIVNLNLSIITTYTETKMDHFIKSKESQLLLKNDKNKYNKSKFCNDLKNFLDYGH 100

Qy 240 SFIGDDMFGNTDRVKGINTKFSDDYKKNVKNLNIKEWKNKKNANLWNIHVK 299
Db 101 LAMGNDMFGGYSTKAENKIQEVFGAGHSEIHEIKNFRKKNWNEFREKLEWMLSEHK 160

Qy 300 GNISKECAIIPABEPQINLMIKEWNEFLMEKRLFLNIKDKCVENKKEACFGGCRIPC 359
Db 161 NNIN-NCKNIPQEBELQITQIKWGHGFLLEDRNRSKLPKSKCKNNTLYEACEKIDPC 219

Qy 360 SSYTSFMKSKTQMEVLTNLYKKNSGVQDNFLNLFK-KNNKD--LDDFFKN-EKEY 415
Db 220 MKYRDWIRSFHWHTLSKEYETQK--VSKENAENYLIKISENKDANKVSLNLCDAEY 277

Qy 416 DDLCDCRTATIKSFLNGPAKNDVDIASQINVDLGRGFCNYKS--NNEKSWNCTGTFT 473
Db 278 SKYCDCHHTTLVKSVLNGNDNTIKKREHLDLDDFSKGCDCNKSVDNTNTKVECKPKY 337

Qy 474 NKFGTCEPPRRQTLCLGRYTLHHRGHEEDYKEHLLGASIEAQLLYKYKBDENALCS 533

```

Db 338 LSTKDVCPPRQELCLGNIDRIYDKNLLMIKEHILAIAIYESRILKRYKNKDDKEVCK 397

Qy 534 IIQNSYADLADIIGSDIIKDYYGKQMEENLNKYNKDKKRNESLKIIFREKMWDEKNENY 593

Db 398 IINKTFADIRDIIGGTDYWNDSLNRKLVGKINTNSKYVHRNKKNDKLFREDWVKVIKKDV 457

Qy 594 WKVMSAVLKNKCTCKDYDKFKIQFLRWFKEWGGDDFCEKKEKIKIYSFSPKVECKKDC 653

Db 458 WNVISWVFKDRTVCKE--DDIENIPQFFRWFSEWGGDYCQDKTMI---ETLKVECKEPC 513

Qy 654 DENTCKNKCSEYKKWIDLKKEYEKQV--DKYTKDNKKMYDNIDEVKNKEANVYLKEK 710

Db 514 EDDNCKSKCNSYKEWISKKKEEYNKQAKQOEYQKGNVYKMY---SEFKSIKPEVYLKY 570

Qy 711 SKECKOVNFDCKIYNESFNEYEDMCKKDEIKYLNIEIKYPKTKHDIYDIDTFSDTFGDGT 770

Db 571 SEKCSNLNFEDEFKEELHSDYKNKCTMCPEVK-----DV 604

Qy 771 PISINANINEQOS 783

Db 605 PISIIRN-NEQTS 616

Search completed: November 21, 2005, 20:38:44
Job time : 274 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2005, 20:26:57 ; Search time 51 seconds
(without alignments)
2282.790 Million cell updates/sec

Title: US-10-677-980-2

Perfect score: 6481

Sequence: 1 MKGYFNIYFLIPLIFLYNVI.....VQETNISDYSEYNNEKNMY 1210

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 80.*

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1557	24.0	1435	2 A37793	erythrocyte-bindin
2	805.5	12.4	1045	2 T18373	erythrocyte bindin
3	765.5	11.8	1070	2 T30848	Duffy receptor - P
4	723.5	11.2	1153	2 T28652	erythrocyte bindin
5	655	10.1	778	2 A35970	erythrocyte-bindin
6	623.5	9.6	2706	2 T28155	variant-specific s
7	600	9.3	3006	2 T28625	variant-specific s
8	555.5	8.6	2924	2 T18378	variant-specific s
9	545.5	8.4	3026	2 T28431	variant surface pr
10	498.5	7.7	3078	2 T28432	variant-specific s
11	483	7.5	2212	2 T28157	erythrocyte membra
12	456	7.0	1701	2 T09127	probable erythrocy
13	438	6.8	1711	2 C71625	variant-specific s
14	433.5	6.7	2664	2 T28636	variant-specific s
15	410.5	6.3	2485	1 H71621	serine/threonine-s
16	409.5	6.3	2647	2 T28161	hypothetical prote
17	401	6.2	3724	2 T18427	hypothetical prote
18	397.5	6.1	2135	2 T14602	variant-specific s
19	395	6.1	2523	2 T18477	hypothetical prote
20	392	6.0	2197	2 B71600	variant-specific s
21	379.5	5.9	1729	2 T18396	erythrocyte membra
22	370.5	5.7	1979	2 C71622	hypothetical prote
23	370.5	5.7	2228	2 T14029	variant-specific s
24	366.5	5.7	1127	2 T28317	ORF MSV156 hypothe
25	362.5	5.6	1516	2 E71619	RAD2 endonuclease
26	359.5	5.5	2401	2 T28676	rhoptry protein -
27	358	5.5	1308	2 E71622	probable membrane
28	357	5.5	3844	2 T18402	asparagine/asparta
29	356.5	5.5	3394	2 T18501	hypothetical prote

ALIGNMENTS

RESULT 1

A37793

erythrocyte-binding antigen 175 - malaria parasite (Plasmodium falciparum)

C;Species: Plasmodium falciparum

C;Date: 30-Apr-1991 #sequence_revision 30-Apr-1991 #text_change 09-Jul-2004

C;Accession: A37793; S11561

R;Sim, B.K.L.; Orlandi, P.A.; Haynes, J.D.; Klotz, F.W.; Carter, J.M.; Camus, D.; Zegans

J. Cell Biol. 111, 1877-1884, 1990

A;Title: Primary structure of the 175K Plasmodium falciparum erythrocyte binding antigen

A;Reference number: A37793; MUID:91035602; PMID:2229177

A;Accession: A37793

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1435 <SIM>

A;Cross-references: UNIPROT:P19214; UNIPARC:UPI0000129B35; GB:X52524; NID:99884; PID:998

R;Sim, B.K.L.

Mol. Biochem. Parasitol. 41, 293-295, 1990

A;Title: Sequence conservation of a functional domain of erythrocyte binding antigen 175

A;Reference number: S11561; MUID:90377299; PMID:2204855

A;Accession: S11561

A;Molecule type: DNA

A;Residues: 995-1158 <SIW>

A;Cross-references: UNIPARC:UPI000017B5PB; EMBL:X52524

Query Match 24.0%; Score 1557; DB 2; Length 1435;

Best Local Similarity 28.5%; Pred. No. 4.5e-59;

Matches 410; Conservative 214; Mismatches 467; Indels 348; Gaps 42;

Qy	1	MKGYNFYELIPLIFLYNVRINESIIIGRTLYNRQDESSDISRVNSPELNNHKTNIYDS	60
Db	1	MKNISIIYFASFVLYFARNEYDI-----KENEFLDVYKEKFNELDKKYGKVNQKT	55
Qy	61	DYEDVNNKLINSFVENKSVKKGSLSPINN-----KTKSY---DIIPPSYSRNDKFNSL	112
Db	56	D-----KKIFTFIENK-----LDILNKKFNKRWKSYGTPDNIDKNWSLIN-----	96
Qy	113	SEHEDSGNTNSNNFANTSISIGKONKQ-----YTFIQKQTH-----	150
Db	97	KHNEEFNNYQSFLSTS--SLIKQNKYVPINAVRVSRILSFLDSRINNGRNTSSNEV	154
Qy	151	LFACGIRKRSIKVICRENSKIT--VCVPDRKIQLCVANFLNSRLETWEKPEIFLISVNT	209
Db	155	LSNCREKRGKMDCKKNDKRNRYVCIIPDRRIQICIVNLIIKIITYTKETWKDHFIEASK	214
Qy	210	EAKLLYNKEGKDPISIFCNELRNSFSDRSFIDGDMDFGNTDRVKGYNITKFSDYKE	269
Db	215	ESQLLAKNDKYNKSKFCNDLKNISFLDYCHLWGNMDMDFGYSYTKAENKIQEVKGAHGE	274
Qy	270	KNVEKLNIIKEWKEKANKLNNHMIIVNHKGNISKECAIIPAESFQNLWIKENWENFLM	329
Db	275	ISEHKIKNFKREWNNEFKLEWEAMLSSEKNNIN--NCKNIPQEBELQITOWIKWHGEPLL	333

```
QY 330 EKKELFLNLIKDKCVENKYYACFGGCLPCSSYTSFMKKSKTQMEVLNLYKKNSGVDK 389
Db 334 ERDNRSKLPSKSKCKNTLYEACEKCDPCMKYRDWIRSKFEWHITLSKEYETQK--VPK 391
QY 390 NNFLNDLFLK--KNNKND--LDQFFKN--BKEVYDLDCCRYTATIIKSPFLNGPAKNDVDIASQ 445
Db 392 ENAENYLIKISEKNDKAVSLLNNCDABSVSKYCDCKGHTTLVKSVLNGNDNTIKEKREH 451
QY 446 INVNDLFGGNYKS--NNEKSNWCTGTFNNKFGTCPPRRQTLCLGRYTLHLRGGHEED 503
Db 452 IDLDDFSKFGCDKNSVDNTNKWECKNPYILSTKDCVPPRRQELCLGNIDRIYDKNLNM 511
QY 504 YKEHLGASIEAQLLYKKEKDNALCSIIQNSYADLADIIGSDIIKDYKQKMEEN 563
Db 512 IKEHILAIATYESILKRYKNKDDKEVCKIINKTFADIRDIIGGTGYWYNDLSNRKLGVK 571
QY 564 LNKVNKDKKNEESLKFIREKWDENKENYKVMASAVLKNKCTCKDYDFKQKIPQFLRWF 623
Db 572 INTNSKYVHRKNDKLFREBWKVVKDWNVSWFKTKVCKE--DDIENIPQFFRWF 630
QY 624 KEMGDDPCERKKEKIYSFESFKVECKKCDENTCKNCKSEYKWKIDLKSEYEQV--- 680
Db 631 SEWGDDYQDQTKMI---ETLKVECKEPCEDDCKSKNSYKESWISKKEEYNKQAKQY 687
QY 681 DKYTKDNKKMYDNIDEVKKEANVYLKESKECKDVNFDKIFNESPNEYEDCKKCD 740
Db 688 OEYQKGNKYMY---SEFKSIKPEVLYLKYSKCSNLNFEDEFKELHSDYKKNCTMCPE 744
QY 741 IKYL-----NE----- 746
Db 745 VKDVPISIRNNEQTSQEAPEENTEIAHRTETPSISEGPKGNEQKERRDDSLSKISVSP 804
QY 747 -----IKPKT----- 752
Db 805 ENSRPETDAKOTSNLLKLGVDISMPKAVIGSSPNDINNVTEQDNI SGVNSKPLSDDV 864
QY 753 ---KHDYID- ---TFSDTFDGTPIGI----- 774
Db 865 RPKKLELDQNSDESETVVNHISKSPSINNDDSGSGSATVSESSSNTGLSIDDDRNG 924
QY 775 -----NANINEQSGKOTSGTNGSETSDSP---VSHEPESDAAINV 812
Db 925 DTFVRTQDANTEVDIRKENADQDEKGADEERHSTSELSPEEKMLDNEGGSNLH 984
QY 813 EKL-----SGDESSSETGILIDN-----DPSVTNNVNEVH-----DAS 846
Db 985 EEVKEHTSNSDNVQSGGI VNMVNEKELKDTLENPSSSLDEGRAHELSBPNLSSDQDMS 1044
QY 847 NTQGSVNTSDITN---GHSESSLNR-----TTNAQDIKIGSGNEQSDN-----QENS 892
Db 1045 NTPGLDNTSEETTERISNNEYKVNEREDERTLTKEYEDIVLKSHMNRSDGDELVDENS 1104
QY 893 SHSNDNSGLTI--GOVPESDNTQNTY-----DSQNPHRDTPNALA-----SLP- 934
Db 1105 DLSIVNDESEDAEAKMGKNTSESHNSQHIESDQKQDMKVTGDLCTTHVQNEISVPV 1164
QY 935 ---SDDKINE-----IEGFDSDRDSNGRGDTTSNTHVYR-----R 967
Db 1165 TGEIDEKLESKESKIHKAEBERLSHTDIHKINPEDRNSNTLHLKDIRNEENERHLTNQN 1224
QY 968 TNIVSERVNSHDFIRNGMANNAHQYIITQ-----IENGLIIRGQESAGSNVNYKDN- 1021
Db 1225 INISQERDLQKHGF---HTWNLLHGDVGSERSQIINHSHGNRQDRGNSGNYLNMRSNN 1280
QY 1022 -----PKRNFSSNDHKKNIQBYNSRDTKRVREEIILKSQNKCNNEYSMEYCTYSDE 1075
Db 1281 NNFNNIFSRNL---YDKLDDLLENENDSTTKELIKLAEINKCNEISVKYCDMH 1337
QY 1076 RNSSPGCSREERKLCQISDYCLKYFNFYISYIYNCIKSEIKSPYKCFKSEGOSSI 1134
Db 1338 EEIPLKCTCKEKNLNCVADYCMYSYFTYDSBEYNTCKREFDDPSYTCFRKEAFSSM 1396
```

RESULT 2

```
T18373
erythrocyte binding protein - Plasmodium cynomolgi
C;Species: Plasmodium cynomolgi
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T18373
R;Okenu, D.M.; Malhotra, P.; Lalitha, P.V.; Chitnis, C.E.; Chauhan, V.S.
Mol. Biochem. Parasitol. 89, 301-306, 1997
A;Title: Cloning and sequence analysis of a gene encoding an erythrocyte binding protein
A;Reference number: Z18923; MUID:98030256; PMID:9364974
A;Accession: T18373
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1045 <OK>
A;Cross-references: UNIPROT:O15852; UNIPARC:UPI000007F49A; EMBL:Y11396; NID:e1042651; P:
C;Genetics:
A;Gene: EBP
A;Introns: 19/3; 983/3; 1008/3; 1032/3

Query Match 12.4%; Score 805.5; DB 2; Length 1045;
Best Local Similarity 26.2%; Pred. NO. 3.4e-27;
Matches 269; Conservative 162; Mismatches 373; Indels 223; Gaps 40;

QY 312 EEPQINLWIKENWENFLMEKKRLFLNIKDK--CVENK--KYEACFGGCLPCSSYTSFMKK 368
Db 56 EEDNIERWLQGTNE-----RRNEENIKYKDYVTEQIKY-----AQMNEKRTSRLKK 103
QY 369 SKTQMEVLNLYKKNSGVDKQNNFLNDLFKNNKNDLDDPFKNKEKEYDDLDCCRYTATII 428
Db 104 SISQNI RNNVREKDGHEKTDKTDNWERCNVLVMDYDTSNGHP-----AWTLDNV 158
QY 429 KSFL-----NGPAKNDVD---IASQ- INVNDLRGF---GCNYSKNNE- KSNW 467
Db 159 LEFVAEGBEHLANSKSGGPNYDIDHKRTISSDVINHAFQLQNVYMKKCNDRKRIRDMW 218
QY 468 CTGTFNNKFGTCPPRRQTLCLGRYTLHLRGGHEEDYKEHLGASIEY-----EA 516
Db 219 CP---TKK--DVCIPDRRYQLCMKEITNLVDNTNTKPHSDITPRKLYLKIYDVTAEG 273
QY 517 QLLKYYKKEKDNALCSIIQNSYADLADIIGSDIIKDYKQKMEENLKNVNDKKNNEE 576
Db 274 NLLLRKNNIYDKLNDLRWSFEDFGDLIMGTDMESIGYSKVKENLRRI-----FGTGQ 329
QY 577 SLKIFREKWDENKENVYKMGSAVLKNNK-----ETCKDYDFKQKIPQFLRWPFKMGDDF 630
Db 330 NAQLHRKQWNEYKEDIWRENMSSVKKLKGNFWI CEKDPANVNEPQIYRWIREWGRDY 389
QY 631 CEKREKESYSPESFKVECKK-----KQCDENTCKNCKSEYKWKIDLKSEYKQVDKY 683
Db 390 ISELPKEL---QKLKEKCDRKFFYYTDIKVCTVLPCKNACILYDQWITRKKQKQWDLVLSKF 446
QY 684 TKDNKKMYDNDIVENKKEANVYLKESKECKDVNFDKIFNESPNEYEDMC----- 735
Db 447 ---KSVKKYQNTETADIVTAYDILKQELDGFNEVAFENEI--NORDNAYIDLCTVTKKNT 502
QY 736 -KKCDIEIKYLINEIKYPTKRHDYIDTFSD--TFGDTPTISINANINEQSGKDTSNIGN 792
Db 503 QKVVTWVNAEASKAPISKAKIQAVDSKKEKKVQGSAGHGNINSAGHNSITGKAVTNGGQ 562
QY 793 S-----ETSDSPSVSHPESDAAINVE-----KLSGDES 820
Db 563 NGNQTPKSNVQRSDIPESASAKNVDPQKYVSERRDDTTSVTSIAEAGKNEGLTNGRPS 622
QY 821 SSET-----RGILDINDPSVTNNVNEVHDA-----SN 847
Db 623 ESTVEAYSAGDGTVNSAYIPVKNSENSLVTHKGLGPKSKDSDNNGSTESKESWAPDSN 682
QY 848 TQGSV-----SNTSDITN---GHSESSLNRTTNAQDIKIGRSGNEQSDNQ 889
Db 683 SKGETEMRODNGKAKATKDSNTSDNTSATTGATGTVDRNINK---SVPEDGNKIYVSK 739
QY 890 ENSHSSDNGSGSLTIGQVPESDNTQNTYDSQNPHRDTPNALASLPDSDKLINEIEGPDSSR 949
```

Db 740 ENEBSSVNGKATGATVGGTNTDRNTENDTENN-----LPAPDN-KQSEGATPLS 787
Qy 950 DSE-----NGRDTTSN--THDVRTNIVSRRVNSHDPFIRNGMAN-----NNAHH 993
Db 788 KTELELNERVHRTNTDTLSKKNKGSGEKDLQKHDFENNMLNABPNYQOTTTDAEGHH 847
Qy 994 QYTOIENNGIIRGOESAGNSVNYKONPKRSNFSSE-----NDHKONTIOEYNSRDTKRVRE 1050
Db 848 R--DSIQND--KGWRKHNKGAFTKPNPNHNLNHNHNLNNGKLDIKEYKYRDVNATRE 902
Qy 1051 EILIKLSONKNSEYCTYSYDERNSPPCPREERKKLCCOISDYCLKYFNFYIEY 1110
Db 903 KLIYSEVRKKNNSISLNYCNSVEDKISS-NTCSREKSNLCCSISDFCLNYPVYSIEY 961
Qy 1111 YNCIKSIKSPYKCFKSEGOSSIPYPAAGILVIVILLSSASRMGKSNEEYDIGESNI 1170
Db 962 HNCMKKEFDPSYKCF-AKGSFRKAYPAAGALLILLILLITSRN-MINKNSEATFNE--- 1016
Qy 1171 EATFEEN 1177
Db 1017 ---FEEH 1020

RESULT 3
T30848
Duffy receptor - Plasmodium vivax
C;Species: Plasmodium vivax
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C;Accession: T30848
R;Fang, X.D.; Kaslow, D.C.; Adams, J.H.; Miller, L.H.
Mol. Biochem. Parasitol. 44, 125-132, 1991
A;Title: Cloning of the Plasmodium vivax Duffy receptor.
A;Reference number: Z20901; MUID:91187056; PMID:1849231
A;Accession: T30848
A;Status: preliminary; translated from GB/EMBL/DBDJB
A;Molecule type: DNA
A;Residues: 1-1070 <P>
A;Cross-references: UNIPROT:P22290; UNIPARC:UPI0000132B95; EMBL:M61095; NID:g160275; PID
C;Genetics:
A;Introns: 19/3; 1006/1; 1032/2; 1057/1
A;Note: PVDR

Query Match 11.8%; Score 765.5; DB 2; Length 1070;
Best Local Similarity 25.0%; Pred. No. 1.8e-25;
Matches 283; Conservative 166; Mismatches 400; Indels 285; Gaps 52;

Qy 272 VEKLNKIKWEKKNKANLWNHMLVNHKG-NISK--ECAIIPAREPOINLIWEKWNFL 328
Db 17 LHKYNNVLLERTITLLECKNEYVKGNGYKLAHGHC---VEEDNLERWLOQTNER-- 70

Qy 329 MEKKRLFLNIKDKVENKYEACPGGCLPCS-SYTSFMKSKTQMEVL-----TNLY 380
Db 71 -----RSEENIKYK--YGVTELKIKYAQMNGKRSSRIKESIYGAHNFNGNSY 116

Qy 381 KKGKNGVDKNFLNDLPKQKNKNDLDDFFKNEKEYDLDLCRY-----TATIIKS 430
Db 117 MEGKDGDKTG-----BEKGEHKTDSTKNGKANLVMLDYETSSNGQPGATLDNVLE 171

Qy 431 FLNGPAKN-----DVD-----IASQINVDLRGF-----GCNYSK-SNNKSW 466
Db 172 FVTGHEGNSRKNSSNGPNYDIDHKTTISSAI-IN--HAFQNTVMKNCNKKRRBRDW 228

Qy 467 NCTGTFNKPFGTCPEPRQTLCLGRYTLHRGHEDY-----KEHLGASTYEA--- 516
Db 229 DC-----NTKKDVCIPRRYQLCMKELTNLVNNTDTNPHRDITPRKLYLKRKLYDAVE 283

Qy 517 -----QLLKYKYEKEDNALCSIIQNSYADLADIIGSDIIKDYKGGKMEENLKV-NKD 570
Db 284 GDLLLLKLNRY-NKD---FCKDIRSLGDFGDIIMGTFMEGIGYSKVVNNLSIRIGTD 339

Qy 571 KGRNESLKIIFREKWDENKENVKMSAVLNK-----ETCKDYDKFKIIPQLRWFK 624
Db 340 EKAQOR-----RKQWNESKAQIWTAMWYSVKRKGKNTWICKLVAVNIEPQIYRWIR 394

Qy 625 EWGDDFCERKEKTIYSPESPVKVC-----KKKDCDENTCKNKCSEYKWKWIDLKSEYE 677
Db 395 EWGRDYVELPTEV---QKLKERDGIKINTDKKVKVPCQACKYDOWITRKKNQWD 451
Qy 678 KQVDKYTKDKKKMYNIDIVNKKEANV-----YLKEKSECKDVNPDKIFNESPNEYE 732
Db 452 VLSNKFISVKNAB-----KVQTAGIVTPYDILKQELDEFNEVAPENET-NKRDCGAYI 502
Qy 733 DMCKKCDIBIKYLNKIEKYPKTHDIYDIDTDSDFG-DGTPI-----GINA 776
Db 503 ELC-----VCSVEAK-KNTQEVVTVNDNAKSOATNSNFIQSPVDSKAEKVPDSTHG 556
Qy 777 NINEQSGKDTSTNGSETSD-----SPVSHE-----PESDAAINVE-----KLSGD 818
Db 557 NVN---SGQDSSTIGRAVTDGQNGQCTPAESDVQSRSDIAESVSANVDPQKSVKESDD 613
Qy 819 -----ESSSETRGILDINDPSVTNNVNEVHDASNTQGSV----- 852
Db 614 TASVTGIAEAGKENLGASNSRPSSESTVEANS PGDDTVNSASIPVSGENPLVTPYNGLRH 673
Qy 853 SNTSDITNGHSESLNRTTNAQDIKIRSG-----NEQSDNQENSHSSDSSGSLT----- 903
Db 674 SKDNSDSDGPAESMANPDSNSK---GETGKGQNDMAKATKDSNSSSDGTSSATGDTTD 729
Qy 904 -----IGQVPSSEDNTQNTYDSQPHRDTPNALASPSDDKINEIEGPDSSRD----- 950
Db 730 AVDREINKGVPEDRDKTVSGDGGEDNSANKDAATVVGEDRIENSAGSSTNDRKNDT 789
Qy 951 SENG-----RGDTTSN--THDVRTNIVSRRVNSHDPFIR 983
Db 790 EKNGASTPDSKQSEDATALSKTLESSTESGDRTTNDTNTSLNKGNGKQKQKQKDFKS 849
Qy 984 NGMANNNAHQYIT-----QIENNGIIRGOESAGNSVNYKONPKRSNFSSE--D 1032
Db 850 NDTNPEEPSDQTTDAEGHRRDSIKND---KAERRKHMKNDTFTKNTNSHHLNNSNNLSN 906

Qy 1033 HKKNIOBYNSRDKRVREIILKLSKQKNKNEYSMEYCTYSYDERNSPPGCSREERKKLC 1092
Db 907 GKLDIKEYKYRDVKATREDIILMSSVRKCNNTISLEYCNSVEDKISS-NTCSREKSNLC 965

Qy 1093 CQISDYCLKYFNFYISYIYCNKIEKSPYKCFKSEGOSSIPYFAAGILWIVLVLSS 1152
Db 966 CSISDFCLNYFDVYSEYLSCKMKPEPSPYKFTKGPKDKTYFAAAGALLILLILLIAS 1025

Qy 1153 ASRWGKSNEEYDIGESNIEATFEE-NNYLNKLSRI-----FNQEVQSTNIDSYS 1200
Db 1026 -RKMKNDSE-----EATFNEPEEYCDNIHRIPLMPNPIEHMQPSTPLDYS 1070

RESULT 4
T28652
erythrocyte binding protein - Plasmodium knowlesi
C;Species: Plasmodium knowlesi
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T28652
R;Adams, J.H.; SIm, B.K.; Dolan, S.A.; Fang, X.; Kaslow, D.C.; Miller, L.H.
Proc. Natl. Acad. Sci. U.S.A. 89, 7085-7089, 1992
A;Title: A family of erythrocyte binding proteins of malaria parasites.
A;Reference number: Z20495; MUID:92357776; PMID:1496004
A;Accession: T28652
A;Status: preliminary; translated from GB/EMBL/DBDJB
A;Molecule type: DNA
A;Residues: 1-1153 <A>
A;Cross-references: UNIPROT:P50493; UNIPARC:UPI0000132B93; EMBL:M90694; NID:g160290; PID
C;Genetics:
A;Introns: 20/3; 1087/3; 1114/2

Query Match 11.2%; Score 723.5; DB 2; Length 1153;
Best Local Similarity 23.6%; Pred. No. 1.2e-23;
Matches 279; Conservative 171; Mismatches 390; Indels 341; Gaps 51;

Qy 253 DRVKGYNITKPF-SDYYKKNVKNLNNIKKEWKEKNKANLWNHMLVNHKGNISKECALIPA 311

Db 27 ERMKGIILLSCENEYVKNNGYKL-----ATGHHYM----- 57
Qy 312 EEPQINLIWIKEMWENFLMEKRLFLNLIKDKC-VE--NKKTEACFGGCRLPSCSSYTFPMKK 368
Db 58 DNDQIERWLQGTDRS---RRVKIEENVKYKYNVEELNTKYEQTKG-----KRINRLIKE 108
Qy 369 SKTQMEVLNLYKKKSGVGNKPNFLNDL-----PKGNKNDLDDFFKRVKEBYD-DLDCDRCY 423
Db 109 S-----TYEAQN--VADNNYIDDKANGVEYKTDNKTNGEGARNVMVLDYDISGSH 157
Qy 424 TATIIK-----SELANGAKNDVDIASQINVNDLRGFG-----CNY 458
Db 158 PDGLINDVNLGTDEGENFLNSKGG-DHPYRNKRKRCMCGVINQTFLOKNVMRRCNN 216
Qy 459 K-SNNEKSMNCTGFTFNKFPCTEPPRRQTLCLGR-TYLLH-----RGHEE-----DYKE 506
Db 217 KRKRGTRDWDCP---TKK--DVCIPDRYQLCMKELTNLVNNTKTHSHNDITFLKLNLIKE 271
Qy 507 HLLGASIYEAQLLYKKEKENALCSIIQNSVADLADIIGSDIIKDYKYGKMEENLNK 566
Db 272 KLTVDAAVEGDLILLKNTNNVYSEDLCDKIKWSLEDFGDIIMGTDMEGIGYSQVVENNLR 331
Qy 567 VNKKKNEESLKIIFREKKWENKENVKNVMSAVLNK-----ETCKDYDKFKQKIPQFL 620
Db 332 VFGTGTKTQ-----LDRKKWNESKYYIWEATILSVKKKLNGYSAWNCKEDVQINVEPIY 387
Qy 621 RWFKEWDDDFCE---KKEKIYSPESPKECKKK-----DCDENTCKNKCSEYKKWID 670
Db 388 RWAREGMDYSELPEQRKI-----KECKDRKLYTNLRICTMSPCNDSCKLYQWIT 441
Qy 671 LKSEYKQVDKTKDKQKMYKQNIDEVKN-KRANVYLKEKSKCKDVNFDKIFNESPN 729
Db 442 RKQKQMDVLSTKFSVVK-----GQIETENITAYDILKQELNGFNEVMPENEI-NKRDN 496
Qy 730 EYEDMC-----KKCDB-1KYLNEIKYPKTKHDIYDITFSDTFDGTGTFPII-NANIN 779
Db 497 VIIDICLAADENKNTQEHLLKLLKLSAPKLETQRSHSTIQPMSSSGAEKVQGLAHGNIN 556
Qy 780 E-----QOSGKDTNSETSDSPV--SHEPESDAAINVEKLS-GBSSSETRGILD 829
Db 557 DAAVKSITDEAARGDQNGQNTVAESNIKGTDNIEENAAKGVDTKFTVRSADTRGATD 616
Qy 830 IN-----VNEVHDASNTQGSVNTS---DITNGHSESSLNRTTNAQDIKI 878
Db 617 ITETGEKLTNYSGSSEITVKENIPGDGIVKDVSAAVENSENPLETKHKIIFPSKONS 676
Qy 839 -----VNEVHDASNTQGSVNTS---DITNGHSESSLNRTTNAQDIKI 878
Db 677 NSENGSMEFKATSSNPITEAVSSSAEGQVQDSAHRSVNTGRDNSTISAATSDDLGSS 736
Qy 879 GRSGNEQSDNQHSSSD-----NSGSLITIG-----QVPSEDN--TQNTYDSQN- 921
Db 737 GDRKVESLTSIENADDGDPVQGLWNLNDPSVAGGGKSHIKTEENEGSQAEIDGNV 796
Qy 922 -----PHR-----DTPNALASLP-SDDKINE-----IEGPDSSRDSNGRGDT 958
Db 797 DIAEQRTATITEVQPERPDLSDTDN--GNVPRSGNKQNEGATALSAGESLESNEVHKTI 854
Qy 959 TSNTHDVRRTNIVSRRVNSHDFTRNGMANNNA----- 991
Db 855 DNTHGLEKNKGNEKDFQKHDFMNDMLNDQTSDDQTSDDQTSDDQTSDDQTSDDQTS 914
Qy 992 -----HHOYITQIENNGIIRGOESAGNSVNY 1018
Db 915 DQTSDDQISSDQTSDDQTSDDQTSDDQTSDDQTSDDQTSDDQTSDDQTSDDQTS 972
Qy 1019 KDNPKRNPFSSENDHKQIQEYNSRDTKRVREIILKSKQNKCNNEYSMEYCTYSDBNS 1078
Db 973 SNGNELYSHNNLNKLNIDQYEHKATREKILMSEVKNCKNRASLKYCNITIEDML 1032
Qy 1079 SPGPCREERBKLCQIISDYCLAYFNFPYSIETYNCKIKSEIKSPYKCFKSEGOSS---IP 1135

Db 1033 S-STCSRERSKNLCCSISDFCLNLYFELYFYFYFNCMKKEFEDESSYECF-TKGSSTGIGIV 1090
Qy 1136 YFAAGAILVIVILLSSASRMGKSNBEYDIGESNIEATPEE 1176
Db 1091 YFATGGAFILILLFFSKKNVANDYEE-----EATFDE 1123
RESULT 5
A35970
erythrocyte-binding protein - Plasmodium knowlesi
C:Species: Plasmodium knowlesi
C:Date: 14-Dec-1990 #sequence_revision 14-Dec-1990 #text_change 09-Jul-2004
C:Accession: A35970
R:Adams, J.H.; Hudson, D.E.; Torii, M.; Ward, G.E.; Wellem, T.B.; Aikawa, M.; Miller, I.
Cell 63, 141-153, 1990
A:Title: The Duffy receptor family of plasmodium knowlesi is located within the micronem
A:Reference number: A35970; MUID:91004213; PMID:2170017
A:Accession: A35970
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-778 <ADA>
A:Cross-references: UNIPROT:P22545; UNIPARC:UPI000016BFA8; GB:M68518; GB:M37513; NID:616
C:Keywords: transmembrane protein
Query Match 10.1%; Score 655; DB 2; Length 778;
Best Local Similarity 24.2%; Pred. No. 6.4e-21;
Matches 192; Conservative 131; Mismatches 276; Indels 194; Gaps 26;
Qy 532 CSIQNSYADLADIIGSDIIKDYKYGKMEENLNKV-NKDKKNEESLKIIFREKKWENK 590
Db 2 CKDIRGLGDFGDIIMGTDMEGIGYSQVVENNLRQVFGTDEKAKQD-----RKQWNESK 56
Qy 591 ENVKNVMSAVLNK--KE-----TKQYDKFKQKIPQLRWFKEWDDDFC-----EK 633
Db 57 EHLWRAMFESIRSLAKEKFWICKKQDVTLKVPQIYRWIREWGRDYMSKLPKEQGLNEK 116
Qy 634 RKEKIYSPESPKECKKDCDENTCKNKCSEYKQKWLIDLKSEYKQVDKYT--KDNKKM 691
Db 117 CASKLY-YNNMAI-----CMLPLCHDACKSYQDWITRKKQDWDLSTKFPSSVKTKQIG 169
Qy 692 YNIDBEVKNKEANVY--LKEKSEKCDVNFDDKIFNESNEYEDMCKCDEIKYLNKEY 749
Db 170 TENI-----ATAYDILKQELNGFKEATFENEI-NKRDNLNHLCPVVEEARKN---- 217
Qy 750 PKTKHDIYDI-----DTFSDTFDGD--TPISINANINEQSGKDTNTGN 792
Db 218 --TQENVKNVSGVESKAASSNPITEAVKSSGEGKVQEDSAHKSYNKGEKSTNEADP 275
Qy 793 SETSDSPVSHPEPESDAAINV-----EKLSDGESSETRGILDIND----- 832
Db 276 GSQSGAPASRSVDEKAGVPALSGAGQGHDKVPPAPAAATBSAVLHSADKTPNTVTENKEG 335
Qy 833 -----PSVTNNVNEV-----HDASNTQGSVNTSDITNGHSES 865
Db 336 TQWDGAAGDGGKAPGPTVSSDVPFSGGKSGSSTASHALAGENGEVHNGTDT---EPKE 392
Qy 866 SLNRTTNAQDIKIGRSGNEQSDNQENSSHSNSGSLTIQVPSEDNTQNTYDSQNPHRD 925
Db 393 DGEKADPQKDIEV--KGKQDTRDSQGLPHTDERATLGETHMEKDTETAGGSTL--- 446
Qy 926 TPNALASLPD-----DKINE-----IEGPDSSRDSNGRGDTTNTHVRTRTIVSE 973
Db 447 TPEQNVSAVDNNGNVPGSGNKQNEGATALSAGESLKSNEHSVHKTIIDNTTHGLEKNKGNE 506
Qy 974 RVNSHDFTRNGMANNNA----- 991
Db 507 KDFQKHDFMNDMLNDQASSDHTSSDQTSDDHTSSDHTSSDHTSSDHTSSDHTSS 566
Qy 992 -----HHOYITQIENNGIIRGOESAGNSVNYKDNPKRNPFSSENDHKQIQEYNSR 1043
Db 567 DQITIDTEGHR--DNVRNPEIKSSDEMSKGDPMRNSNSELVSHNNLNKLNKLRDQYEHR 624
Qy 1044 DTKRVREIILKSKQNKCNNEYSMEYCTYSDBNSSPGCSREERBKLCQIISDYCLKYF 1103

RESULT 7
T28625
variant-specific surface protein 3 - malaria parasite (Plasmodium falciparum)
C/Species: Plasmodium falciparum
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T28625
R/Su, X.Z.; Heatwole, V.M.; Wertheimer, S.P.; Guinet, F.; Herrfeldt, J.A.; Peterson, D.S.
Cell 82, 89-100, 1995
A/Title: The large diverse gene family var encodes proteins involved in cytoadherence and
A/Reference number: Z20487; MUID:95330813; PMID:7606788
A/Accession: T28625
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-3006 <SUX>
A/Cross-references: UNIPROT.Q26032; UNIPARC:UPI0000078E68; EMBL.L40609; NID:9886376; PID
C/Genetics:
A/Gene: var-3
A/Introns: 2597/3

Query Match 9.3%; Score 600; DB 2; Length 3006;
Best local similarity 19.9%; Pred. No. 6.6e-18;
Matches 318; Conservative 214; Mismatches 493; Indels 570; Gaps 73;

QY 33 NRODESSDI SRVSP ELNNHKTNIYDSYEDVNN-----K LINSFVENKSVKKRS 84
DB 718 NNSNECDVSK-----KVTNPGCNPSASNNLVRVRLAEMMQRYARKOLEKRGGE 769

QY 85 LSFNNKTSYDIIPPSYRN-----DKFNSLSENEGNTNSNFPANTSE- ISIGKON 139
DB 770 INLKGDATKG-----TYRQGPADGPKNVCSINQNHNTVNQNNRAYTYQGPGCKDG 821

QY 140 KQYTFIQKTHLFCAGIKRK- SIKWICRESEKI-----TVCVPDRKIQLCVAN--FLNS 191
DB 822 SN-----GGVMMKIGTPW---KPGROQMSAEDIYMPRRQHMCTSNLEYLQT 866

QY 192 RLETMEK-----FKEIFLISVNTAKL-----LYNKNKGK-----DPSIFCNELRN 232
DB 867 KDGPKLQGDGKLVNNSFLGDLVLSAKMDAGKI IELYKQNNKSNLTPEDNESACRALRY 926

QY 233 SFSDFRSFIDGDMDFGNTDRVRKYIN-----TKFSD-----YYKEKNVEKLNNI 278
DB 927 SFADLGDIIIRGDL-WDKNSDAKRLQTNLKEIFTKIKEELPEDIKKKYDKDGDTHKL--L 983

QY 279 KKEWWEKNANLWNNMIVNHKGNISKECAIIPAE--PQNLWIKENWENFLMEKRLFL 336
DB 984 REDWWEANRHQVWRAKCAIENDKMKCNGIPIEDYIPORLRWMTWAEFCKEQSRLYN 1043

QY 337 NIKDKCVENK-KYBAC-----FGCRLPCSSYTSFMKKSKTQMEVLTNLY-----380
DB 1044 KLVDCKSKGKAKSCCTQKQGDCTKCAACDYNKTKI PWEEQWEKTKNKYALYKKALD 1103

QY 381-----KKNSGVDRN-----NFLNDLFKN-----400
DB 1104 SVNGKESKKKTASDAKDQVVFHFLAELIRKSGGKGGKGNKVTTSPTTPTNTLYSSAAG 1163

QY 401-----NNDLDPFRNEKEYDDLCDCRYTATIIKFLN-----433
DB 1164 YIHHELGRVTGCTQKBFYCSKNGKYAFKDPKGYEACK-----NDRNPKQP 1212

QY 434-GPAKNVDVITASQI-----NVDLRGFCNVKSNNEK--SNWCTGTPTNKPPGCEP 482
DB 1213 PAPKKEDEADCVVKPLLKQKGETDDIDGNCQKRYKAGDKYKPGWDCNSQIHTHTHINGACMP 1272

QY 483 PRRTCLGLRTHLLHRGHEEDY-KEHLIGASIVYEAQLLYKYKEKDNALCSI-----534
DB 1273 PRQKLCVSGJTKTDRIKAIYIRTEFIKSAIETHFAWDRYKEDNGEAEALKNGNIPE 1332

QY 535-----IQNSVADLADIIGSDIHKDY-----GKQWENLNKVNKDK---KRNEESLKIFR 582
DB 1333 GFKEQMYTYTFGDRDIFFGDRDISHAYTSGVSPKVIITILSKENDAKYAAKQNSNELL--1390
QY 583 EKWDENKENVWYMSAVL-----KNKETCKDY-----DKFQKLPQF 619

RESULT 8

T18378

variant-specific surface protein 1 - malaria parasite (Plasmodium falciparum)
N/Alternate names: erythrocyte membrane binding protein 1 (EMP1)

C/Species: Plasmodium falciparum

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C/Accession: T18378

R/Baruch, D.I.; Pasloske, B.L.; Singh, H.B.; Bi, X.; Ma, X.C.; Feldman, M.; Taraschi, T.

Cell 82, 77-87, 1995

A/Title: Cloning the P. falciparum gene encoding PfEMP1, a malarial variant antigen and

A/Reference number: Z18925; MUID:95330812; PMID:7541722

A/Accession: T18378

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

DB 1391 DQWHDQKDIWEGMLCALTHKISDEKKKEIKNKYSYKLLNESPKGSKNKVDFAKKPQF 1450
QY 620 LRWFKEWGDDPCEKREKIKIYSPESFKVECK-KKDCD--ENT-----CKNKCSEYKKWIDL 671
DB 1451 LRWFIEWGDEFCQAOREEK--EAKVKVSCSDAKYDGCNKTKSNASCVSACKVYEDYITK 1507
QY 672 KKSEYEKQVDKTKDK--NKKMYDNIDEVNKEANYVLKESKECKDVNFDKIKFIESPN 729
DB 1508 KKVETTKQKGFDAEKITDKGEGYGF---STKDASEYLK---KKCLD-----1548
QY 730 EYEDMCKKCDKILYNLEIKYPKTKHDYIDITPDSDFDGTPIISINANINEQ-----781
DB 1549 ---DTCNCKQVK--NNTBYWNTPNKTY-----TNSNLEKCECOPPOE 1587
QY 782 ----QSGKDTNMTGNSDTSPPVSHS-----PESDAAINVEKLSGDESSESTR 825
DB 1588 PPGPEGGARSDSGPRDTPRPAGSDARNTVPSPPPRPAGDTVHEVAEVEEEDDGD 1647
QY 826 GILDINDPVTNNVNEVHDASNTQGSVNTSDITNGHSESSLNRTTNAQDIKIGRSGNEQ 885
DB 1648 DLPEDQDEDV-----EVAGAEEDDLVGVARILGR-----TNSPDEDEDEASEE 1694
QY 886 SDNOENSHSSDNGSLTIGOVPS---EDTQNTQYD-----SONPHRDTPNAL 930
DB 1695 DDDDDQADTTTEVGO---GEEETAEDHQDTTEETVDQEKAEEDKGGGETPOKETQPKV 1751
QY 931 ASLPSP-----DK-----I 939
DB 1752 EVNFCDIVKTLFTTTTETLKEACPTKYVNGREKFPNWKCISSGSDASGSSICIPPRRRKLYL 1811
QY 940 NELEGDFSRSDENGSDTTSN--THDVRTNIVSERRVNSHDIFRNGMANNNAHQY--I 996
DB 1812 HKIEGVDTTVSSD---GETTPTITHDALREAFITQAAVETFL-----WHRYKKI 1858
QY 997 TQIENNGIIRGOEE-----SAGNSVNVKNP-----KRSNFSSEND 1032
DB 1859 KEKER---QBELQNGTFLPAPQKVPEDNPEHPQKLEKGIPEEFKROMFTLGD 1912
QY 1033 HK-----KNIOEVNSR-----1043
DB 1913 YRDLGVGVKDDVAQALEASGDNKSGDKNIKIDSEKIKSVIEKSGEQTTPGPKPGQTTTKP 1972
QY 1044 -----DTKRVREI-----IKLSKQNKCNKNSMEYCTVSDE 1076
DB 1973 EEWQKNGEHIWNAITCALHTNTDTRQVDDQKQLFENGKNTPKNSQYQKNVTISSV- 2031
QY 1077 NSSPGPCSRERKKLCCQIS-----DYCLKYVNFVYSIEY-----NCIKS 1116
DB 2032 SNGGPIGNIKLEQFASRPTFLRWLEWGEFPCR--QYHKLRIKEECHKGNRNCDDD 2089
QY 1117 EIKSPYKCFKSGQSSIPYFAAGILVVIVLLSSASRMKGNSREBYDIGESNIEATFEE 1176
DB 2090 GPECKE-MCPCKDGSFE-----TLKPCSCAKSKYKXW-ISRKKDEFTKQK 2134
QY 1177 NNYLNLKSLRIFNOEVOETNISDYSEY-----NTNE 1206
DB 2135 GAYEKQ-----KKDAEGNNNDYKEFSKTLRNYND 2163

RESULT 8

T18378

variant-specific surface protein 1 - malaria parasite (Plasmodium falciparum)
N/Alternate names: erythrocyte membrane binding protein 1 (EMP1)

C/Species: Plasmodium falciparum

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C/Accession: T18378

R/Baruch, D.I.; Pasloske, B.L.; Singh, H.B.; Bi, X.; Ma, X.C.; Feldman, M.; Taraschi, T.

Cell 82, 77-87, 1995

A/Title: Cloning the P. falciparum gene encoding PfEMP1, a malarial variant antigen and

A/Reference number: Z18925; MUID:95330812; PMID:7541722

A/Accession: T18378

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A;Residues: 1-2924 <BAR>
A;Cross-references: UNIPROT:Q25733; UNIPARC:UPI0000079077; EMBL:U27338; NID:g914918; PID:
C;Genetics: A;Gene: EMP1
A;Introns: 2476/3

Query Match 8.6%; Score 555.5; DB 2; Length 2924;
Best Local Similarity 22.7%; Pred. No. 5e-16;
Matches 280; Conservative 182; Mismatches 434; Indels 337; Gaps 65;

Qy	48	ELNNHKTNIYDSYEDVANKLINSFVENKSVKKRSLSFINNKTSYDIIPESYSYRND	107
		: : : :	:
Dd	1674	KINNELKGN-----GKDFNGKCN-----VKONGAVI-----GGSC	1706

Qy	108 KENSLSENEEDNSGNTSNNFANTSEISICKDNKQYTP-IQKRTHLPACGIKRKSIIKWICR 166
Db	1707 KEEOQTYENSVN-----NINNK-----KDNERPKIGQWKNFKYI GTTRKDL----- 1749

Qy	167	ENSEKITVCVPDRKIQLCVANF-----LNSRLTEWKFKEFLISVNTAEKLLYKNNE	219
Db	1750	-----CIPPREHMCDDLSMLGRTTISDSALLKKIQEAAKSERDDIRKLLEQNS	1801

Qy	220	GKDPSICFNELRNFSDFSRSPFGD---DMDFGGNTDRVKGY---	INTKFSDYVK-----268
Db	1802	CDEHRI--CDAMKYSFAD-----LGDITGRDLWNKNCKQGLKRLEAFINYNKLQND	1855

Qy	269	----	E	K	V	E	K	L	N	N	I	K	E	W	E	K	K	N	A	L	N	H	M	I	V	N	----	H	K	----	300																						
			:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:																												
Db	1856	K	N	K	Y	E	R	P	K	L	Q	L	R	S	D	W	D	A	N	E	K	H	I	N	W	A	M	T	C	N	A	P	D	A	K	F	L	K	G	N	P	N	T	S	G	S	S	K	G	I	M	T	1915

	Qy	301	---	NISKECAIPAE	--PQINLTWIKENNENF--	LMEKRLFLNLIKDKVCVEN----	345
	Dz	1916	HSCNGYDKE---	PPDYDIPOPRRWMQWSSEFCLLNEEMEQFKTGCCKNSITCE	1971		

QY 346 -- KKYEACFGGRLPCSSYTSFMKSK ---TQMEVLNLYKKNSGVDKNNFLNDLFFK 399

Dd 1972 DDRNGTNC - ENCKNQCEKYKGLIHWNKLGFDKYKEIYNEIYNNKDSKINSNEYFKKELEK 2030

```

QY      400      -----NNKMDLDUFFKN-EKEYDDLLCDC----- 421
          |||         |||
Db    2031   LKDCKELNSDKCIDBEATHCTKYKFNSENKNHNHYAFKNPKPEYERACKCDAPDPLDN 2090

```

QY 422 --KATATILKSFNLGPKAKNDVJLASQILNVNDLKGFGCNK--SNNEKSWNCIGFIFKNK 4 7 7
Db 2091 CPKDSATYEK-----ACNTL-LPTKL-----CESKTFNDDSDWT--SFVQTSP 2132

QY	4 7 8	---	G	C	E	P	R	R	Q	L	C	K	I	L	L	H	R	G	H	-	-	E	D	I	K	H	L	G	A	S	E	A	L	L	K	I	K	R	E	K	S	N	A	L	5 3 1									
Db	2133		R	D	N	T	G	V	L	P	P	R	R	Q	I	C	L	N	I	T	T	K	R	S	T	E	K	I	D	F	K	A	E	L	M	T	S	A	N	E	G	K	L	C	E	L	Y	-	K	R	D	V	T	2191

D6
2192 LOAMKSPFYDGVIRKGTDLISTAPDLKTLLNVLLKGDGTNE--IKEDRGKWWTENRT 2249

Db 2250 RVWHAMLCGYKAAAGGKTEERDCSLPD--DNTHQFLRWFREWSHEFCAKRQK---LFNEVK 2304

2305 RECASACIIIEYGTIDPPVCEACTQRYDIIITRKIQETRLINYOYNTNFENEKAEVTKAP 2364

2365 EFNDKCNDCNCLSYDIEKKWKNYDSFDNDLKNKICIQIKPK-RPPKKVKEEE 2423

[illegible]

D_b

2473 LFLKKTKSSVGNL---FQILHIPKSDYDIPT-KLSPNRYIPYTSGYRG----- 2518

[illegible]

RESULT 9

T28431
variant surface protein 1 homolog AAVAR - malaria parasite (Plasmodium falciparum) (fragment)
C:Species: Plasmodium falciparum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T28431
R:Smith, J.D.; Kyes, S.; Craig, A.G.; Fagan, T.; Hudson-Taylor, D.; Miller, L.H.; Baruch, A.;
Mol. Biochem. Parasitol. 97, 133-148, 1998
A:Title: Analysis of adhesive domains from the AAVAR Plasmodium falciparum erythrocyte m
A:Reference number: T20486; MUID:99094502; PMID:9879893
A:Accession: T28431
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-3026 <SMI>
A:Cross-references: UNIPROT:Q26030; UNIPARC:UPI000007B327; EMBL:L42244; NID:G3540144; PI
C:Genetics:
A:Gene: var
A:Introns: 2906/3

Query Match 8.4%; Score 545.5; DB 2; Length 3026;
Best Local Similarity 21.7%; Pred. No. 1.4e-15;
Matches 309; Conservative 210; Mismatches 455; Indels 447; Gaps 77;

Qy	28	GRTLYNRQD----	ESSDISRVNS-----	PELNHNKTNIVDSO	61
Dd	1564	GHRIVCSGDGHI CEKTDTSRNNTFIDLHPCRLKECIKYRWRWIEKKKEBFHQNKYNYEKE	:	:	:
Qy	62	YEDVNNKL-----	INSFVENKSVK-----	KBSLSFINNKTS-----	YDIIPPS
Dd	1624	F--NNDLKEGGYSSPNFNFLASLHCKGHNRDDDKNKIEF-NNHTKTFGPSEYCKACPV	:	:	:
Qy	102	YSYRNDRKFNLSB--	NEDNSGNTSNNFANTSISICK-----	DNKOYTFIQKETHLPACG	155
Dd	1680	YGVCNKNGCECETIHKHTDLGNQDNNTYTDIKVLVIDRKGESNDEELKNVCNNTSLF---	:	:	:
Qy	156	IKRKS IK-WTCRENSEKITVCPDRKIQLCVANPLANSRLRTEMTKFEBIFILSVNTEAKLL	:	:	:
Dd	1737	-KDSSVQYWKRCQKNIE-VDOCCIID-----	NFLDIDIKYMEFNVF-----	QRWL R	1780
Qy	215	YNKNEGKDPSIFCNELRNSFSDFRSSFIGDDMPGGTDRVKGVINTKFPSDYYKEKNVE-	:	:	:
Dd	1781	Y-----	FV-----	HYNILKDKIKECIKTK--	DEKSNKCING
Qy	274	---KLANNIKKBWEKONKANLWNHMIVHNHGNISSKECAIIPABEPQINLWIKEW-NENFLM	:	:	:
Dd	1811	CKGKI.ECVKK--	WLKQKODESWNKDHYEKNKSLGYGI PH-----	WVKSYPVEQLY F	1861
Qy	330	EK--KRULFNLIKCVKENKCYEACFGGCRUPCSSYSTSFMMKSKSTQMTEVLNLTYLKONGSV	:	:	:
Dd	1862	DKDYKKAQDVIEDB-NEKKIWGTGJ--	VECTN-----	BETKENKOFITNLKLOEKI	1913

Qy 388 DKNFLN-----DLFKNNKNDLD-----DFFKNEKEYD-----DLCDCRY 423
Db 1914 ESCQTOHNPNGTKPCDEIPPHSDTELDEQDTDTDDNSDKIYDTKPPCPKQVEDTKE 1973
Qy 424 T-----ATIKSFLNGPAKNDVDIASQINVDLGRFGCNVY-----SNNEKSWN 467
Db 1974 TEXPKVLPGPPDACEIVEILNGO-----DGTKKIE-----ECNTKYPTKNDYPGWN 2021
Qy 468 CTGTFNFKPGTCEPPRRQTLCL-GRTYLLHRGHEEDYKEHLLGASIEYEAQLLYKYKE- 525
Db 2022 CTDKVINRESGSCMPRRQKLCIHNLEHLSKATETELRKAFTICAAIETFWLWDYKED 2081
Qy 526 -KDENAL-----CSIIQN-----SYADLADIKSGDIKDY----- 555
Db 2082 KDEKTEGGISDDPDPPOKKLEGGTIPEDFKRQMFYTYG DYRDLFGTDISKGHGKES 2141
Qy 556 -YGKQMBENLNKYNKOKKRNEESLKI PREKWBENKENWKMVSAVLK-----NKETCK 608
Db 2142 ALGKKI-DSLFLK-NGDQSPSGKTP-----TEWWDYGDPIWKGWVCGLSHHIKNGNKEQLR 2196
Qy 609 -----DYDKFOKI-----PQFLRWFKWGGDDFCEKKEKIKIYGFESPKVECKKDC- 653
Db 2197 KNLTDNNKYTKISKLEDFASRQFLRWFTIEWGQDFC---RERVVKINQLKTCGNEYECG 2253
Qy 654 -DEN-----TCNKCSEYKWKIDLKSEYKQVDKYTKOKKKMYD-----NIDEVKNKEAN 704
Db 2254 SQENGKKEACKNACEAYKSLWKMDQOYEQOTAKFKDKKDKFDTGSAEVDVAAVSSVH 2313
Qy 705 VYLKESK-----ECKDWNFDDKLFNESPNEYEDMCK----- 736
Db 2314 EYLQEBELKNLCTKDCACMEKPSAQDEBETELLGNGYFPEAWDPPKIEGERCKALPSEP 2373
Qy 737 -KDE--IKVLYNE-----IKY-----PKTKHDIYDIDTSDTFDGTPIISINANIN 779
Db 2374 MSCVEQIAKLHREAKENKVIYESSLKGTPAKSKNDCTKID--EAIKGDNGSKIINKSI- 2430
Qy 780 EQSGKDTSTNGSETSDSPVSEHPESDAAINVEKLSGDESS---SETRGILINDPSVT 836
Db 2431 -----LDSTFASNCEQSEKDATRLKIGKQWQFNKINGTETKLYVPPRRKDMCFND---L 2482
Qy 837 NNV--NEVHDASNTQGSVNTS-----DI--TNGHSESLNRTTNAQ-----DIKI 878
Db 2483 KNTQFNEVQNSLLEKIQHVAKNEGIDILKLNPDQNAFSELCAMKYSFADLGDILR 2542
Qy 879 GRSGNEQSDN-----QENSSHSDNSGSLTIGQVPESDNTQNTYDSQNPHRDTPNALA 931
Db 2543 GRSKIDPTNNKIEKELQKIFKQIODDNASLSKMELP--ELREKWDAN--RKEVWNAMT 2598
Qy 932 SL-PSDDKINEIEGDFSSRDSNGRGDTT-----SNTHDVRTNIVSERRVNSHDFIRNGM 986
Db 2599 CVAPND AHL-----KKKQNNPGNKSQIIASQTEQTKK-----SHD----- 2634
Qy 987 ANNAHQQYITQIENNGIIRGQESAGNSVNYKDNPKRSNFFSENDDHKNIQEYNSRDTK 1046
Db 2635 -SEPPDYDI-----PERYF-----LQWSEYCYCK 2659
Qy 1047 RVREIILKSKQNKCNNEYNMEYCTYSDERNSSPGCSREERKKLCCQISDYCLKYPNF- 1105
Db 2660 ALKE-----KNDEMKNDCSK--CI-----KSGATCEKEEDKECKECNCKKEYKNIV 2705
Qy 1106 --YSIEYNNCKISEIKSPEYKCPKSEGOSSIPYFAAGGILVIVILLSSASRMKSNEEY 1163
Db 2706 DKWQSEF-----DQOQLYKLYQTDRTGHP-----STARRNPSTIEFT 2743
Qy 1164 DIGESNTEATFEENYLNKLSRIFNOBVOETNISDYSEYNY 1204
Db 2744 QKLEDS CNDYSDAKYDLIDISTHCTDYKSET-----DSRESNY 2781

RESULT 10
T28432
variant-specific surface protein 1 - malaria parasite (Plasmodium falciparum)
N;Alternate names: erythrocyte membrane binding protein 1 (EMP1)

C;Species: Plasmodium falciparum
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T28432
R;Su, X.Z.; Heatwole, V.M.; Wertheimer, S.P.; Guinet, F.; Herrfeldt, J.A.; Peterson, D.S.
Cell 82, 89-100, 1995
A;Title: The large diverse gene family var encodes proteins involved in cytoadherence and
A;Reference number: Z20487; MUID:95330813; PMID:7606788
A;Accession: T28432
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-3078 <SUX>
A;Cross-references: UNIPROT:Q26031; UNIPARC:UPI00000820C2; EMBL:L40608; NID:9886374; PID
C;Genetics:
A;Gene: var-1
A;Introns: 2611/3

Query Match 7.7%; Score 498.5; DB 2; Length 3078;
Best Local Similarity 18.9%; Pred. No. 1.4e-13;
Matches 317; Conservative 180; Mismatches 520; Indels 661; Gaps 67;

Qy 90 NKTYSYDIIPPSYSYRNDKFNLSLENEDNSGNTNSNPNANTSE-ISIGKDNKOYTFIQKR 148
Db 1309 NKEYTEKQPPPEYATACDCINRSQTEEPKPKERNVESACKIVBKILEGKNG-----R 1360
Qy 149 THLPACGIRKSIKWI CRENSE--KITVCVDPKRIQLCVANFLN-----SRLETMEKEKEI 202
Db 1361 TTVGECNPKESYDPMDCKNIDISHDGACMPRRQKLCIYYIAHESQTENIKRTDDNLKDA 1420
Qy 203 FLISV-----NTEAKLLYNKNEKQPSIFCNELNRNSFSPSRSSFIGDDMD 247
Db 1421 FIKTAAATFLSWQYKYKNDSEAKIL---DRGLIFSQFLSRMMYTFG DYR-----D 1469
Qy 248 FGNTDRVKGYINTKPSDYKKEKNVEKLNLI-----KKEWKEKNKANLNHNM 294
Db 1470 ICLNTD-----ISKQNDVAKAK--DKIGKFFSKGSKSPSGLSRQEWKWTNGPEIWKGM 1522
Qy 295 IV-----NHNKGNISKECALIPABEPQINLIWKEWENFL 328
Db 1523 LCALTKYVTDITDNKRKIKNDYSYDVKNQSGNPSLEEF---AAKPQFLRMWIEWEEFC 1579
Qy 329 MEKKLFLNLIKDCVENKKEYACFGG---CRLPCCSYTSFMKSKTQMEVLTN----- 378
Db 1580 AERQKKEKNIKACNEINSTQCNDAKHRCNQACRAYQEVVENKKEKFFSGQTNFVLKAN 1639
Qy 379 -----LYKKNSGVKNNFLNLDLFKKNKNLDL-----DFFKNEKEYD 416
Db 1640 VQPDPEYKGYEKDGVQPIQGNAYLLQKCDNNKSCMDGNVLSVSPKPEKPFKYAHKYP 1699
Qy 417 DLCDCHYTATIISKFLNGPAKND-----VDIASQINV--NDLRGP-----GCYKSN 461
Db 1700 EKCDC-YQGHKVPISIPPPPPVQPOPEAPTIVTVVDCSIVKTLFKDTNNFSDACGLKYGKT 1758
Qy 462 NEKSNWCTGTFTNKFPGT-----CEPPRQTLCLGR----- 492
Db 1759 APSSWKCIFSDYKSGAGATTKSGSDSGSICIPRRRLYVGLQEWATALPOGEAAPS 1818
Qy 493 -----TYLLHRGHEEDYKEHLLGASIEYEAQLLYKYKEKDNALCSCI 534
Db 1819 HSRADDLRNAFIQSAAIETFFLWDRYKEBKQPQDGSQQALSQLS-TYSDDDEEDPDKL 1877
Qy 535 IQN-----SYADLADIKSGDIIKDYKKEENLNKYNKDKKRNEESLKIF 581
Db 1878 LQNGKTPDPLRMFLMYTLGDYRDLVHGNGTSD-SGNTNGSNNNNIVLEASGNKEDMOKI 1936
Qy 582 RE-----KWDENKENWKMVSAVL-----KNEKTC----- 607
Db 1937 QEKIEQILPKNGGTPLVPKSSAQTPDKWNEHAESIKWGMICALTYTEKKNPDTISARGDEN 1996
Qy 608 ---KD---YDKF-----QKI----- 616
Db 1997 KIEKDEVEKEFGSGTADKHGTASTPTGYTKTDYVEKVKLEDTSGAKTPSSASDTPLLS 2056
Qy 617 ----PQFLRWFKWGGDDFCEKKEKIKIYSPESPKVECKKK----- 651


```
QY 636 EKIYSPSPKVC-----KK-----KDCDE-----NTCKKCYSEYKWI 669
Db 1626 KRL-----EKIVCEWDEGKKQKSGDGECEIRKQDYSTRVDFYCEPCQKCYRFRKWI 1682
QY 670 DLKKSEYKQVDTYKDKNKKMVDNIDEVKN-----KEANVYKBE-KSKECKDVNFDD- 721
Db 1683 EKKDEYDKQKEAYNNQKTARRNNNDNAFSTLDTCTTAGDFLOTLKNGPCNDNVDDSD 1742
QY 722 -----KIPNESPN-EYEDMCKCDEIKYLBNEIKYPTKTHDIYDIDTFSDTFGDTGISIN 775
Db 1743 GENKKIPDENGDTFKYQYCGTCS-----LNGFK-----CNGDDCVRVTN 1782
QY 776 ANINEQO-----SGKDTSNSTGSETSDSPVSEHES-----DAAINVEKLGDRSS-- 821
Db 1783 VTCNGSNRRTTITADDIKNGGNSABINMLVSDDINGNGFNDLEACKNANIFKGIKENKW 1842
QY 822 -----SETRGILDIND-----PSVTNNVEVHDASNTQGSVSTSD 857
Db 1843 KCVYFCKSDVCGLKKNNDIDQNIILIRALFKRWLEFLDYDKIRKLNPCINNGEKAI 1902
QY 858 ITNGHSESSLN-RTTNAQDIKIGRSGNEQSDNQSNSHSDNSGSL-TIGQVPSDNTQN 915
Db 1903 CTNGCVBQWIMQKRTETNIK--NRFNEQYNGDDTEMKSSFRSLVDLIRQIAA----- 1954
QY 916 TYDSQNPRTDTPNALASLPDSDDKINEIEGDFSSRDSSENGR-GDTTSNTHDVRRTNIVSER 974
Db 1955 TIDKGN-----HNGLVKLKVSVKCN-----CGNNSQNGKEGE----- 1986
QY 975 RVNSHDFIRGMANNNAHQVITQIENNGIIRGOESAGNSVNYKONPKESNFSS----- 1029
Db 1987 -----ENDLCLLQKLEKKAECCKONPETSIGIPQOPCEV 2021
QY 1030 -----ENDHKNTQYNSRDTKRVREBIIKLSKQNKCNNEYSMEYCTYSDERNSSPCPC- 1083
Db 2022 SPNHIEDEQPLEEENVEHPKICDDVLKTPQ-----PEEPGETC-----EESPGPTD 2071
QY 1084 ---SREERKK 1090
Db 2072 VKEEEEEKK 2081

RESULT 12
T09127
Probable erythrocyte-binding protein MAEBL - Plasmodium yoelii
C;Species: Plasmodium yoelii
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: T09127
R;Kapke, S.H.I.; Noe, A.R.; Fraser, T.S.; Blair, P.L.; Adams, J.H.
Proc. Natl. Acad. Sci. U.S.A. 95, 1230-1235, 1998
A;Title: A family of chimeric erythrocyte binding proteins of malaria parasites.
A;Reference number: 216577; MUID:98115903; PMID:9448314
A;Accession: T09127
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1701 <KAP>
A;Cross-references: UNIPROT:O61164; UNIPARC:UPI000007D433; EMBL:AF031886; NID:G2947227;
A;Experimental source: subspecies yoelii; strain YM
C;Genetics:
A;Gene: maeb1
C;Keywords: alternative splicing; cell binding; erythrocyte invasion

Query Match 7.0%; Score 456; DB 2; Length 1701;
Best Local Similarity 19.9%; Pred. No. 4.8e-12;
Matches 273; Conservative 192; Mismatches 510; Indels 398; Gaps 57;

QY 33 NRODESSDISRV-----NSPELNNHKTNIYDSYEDVNNKL-----INSFVENKSVKK 81
Db 436 NQNIYEPDIENVKTIWKNRNSLSNLKYNKDKFKNSDINKGMAMNMDINEIKENSGLQT 495
QY 82 KRSLSFNNKTK-----SYDIIPSPSYSRNDK-----FNSLSSENDNS-GNTNS-NNF 127
Db 496 NK-----GNETKTKYGLYNYPTIPISYLQIHHKMKELKYNMDSNSFTSFHNTNAPTHY 550
```

```
QY 128 ANTSEISIGKNNKQYT-----FIQKRTHLFACGI 156
Db 551 EGNKSFSTGVNNKRENTYGTQDINLNRNNYQPKNPNQPBAYMDRFDIEKNHIY--- 606
QY 157 KRKSIKW-----ICRESEKIITVCVDPDRKIQLCVAFNLSRLB-TWEKPK 200
Db 607 ----IDWKQDGKYGSGKLKYNIIISHETADTIOSLLITDKDDICPNHYSQRAQSCPNYG 662
QY 201 EIFLI-----SVNTEAKLLYNKNEGKPSIFCNELRNSFSDFRSSFIGDDMDFGNTDRVK 256
Db 663 KSIIVKTPESINGNEHL--NSN-----FLNEIRT----- 689
QY 257 GYINTKFSYIYKKNVE-----KLNNIKKWEKWN--KANLWNH----- 293
Db 690 GYLNK-----YKSNVELPEYKSLAMHGDLSVCPKSWDEENLYKKNRDYNDMCKSTV 744
QY 294 -----MIVNHK-----GNISKECALIPAEPOINLWIKENWENFLM 329
Db 745 MKSTIPLKMPDYKTKKLYGLYGLGRLGNSISKYKNIFKSPQNNITL-----PM 795
QY 330 EKKRLPLNIKD-----KCVENKKYBACFGGCLPCSSYTSFMKCKSKTOMEVL--- 376
Db 796 FNPSSIKNLLDCLSYCYCLGPCLENNYNNKCP-----RSLPAYFNHETNECIIILGTHE 848
QY 377 ---TNLYKKKNSGVDRKNFNLDFKKN-----NKNDLDDFPKNEKEYDDLCDCRYTATI-- 427
Db 849 QERNNNCRTRSDTDKPNQCN--VRKNISTKNWYVTSFIR--PDYEKCPPRYPLKPKS 904
QY 428 -----IKSFLNGPAKNVDI-----SFTSLTNTDI 1002
Db 905 FGKIDETGCKSLIN--KXNIINIPFLSCLEYMFIMYPSVLQRTKKNYNGWVWVASES 962
QY 443 ASQINVNDLURGFCNYSNNKSNWCTGTFTNKPFGTCPPRRQTLCLGRTYLLHKGHE 502
Db 963 VNSSNLNNAKG-ECYY--INEXP-NCVIDKNHF----- 1002
QY 503 DYKEHLLGASIVEAQL-----LKYKYEKDENALCSIIQ-----NSYA 540
Db 1003 DFNQINLVKLDELVINNDQSSSHNRAKYNTPIENSESTIVRKHNSAPEHFRSLKINSYT 1062
QY 541 DL---ADIITGSDI IKDYGKMEENLNKVKDKGNEESLKIIFREKW---WDEKENVW 594
Db 1063 PNRGENFAKESDSTRTNTDESXMDVIRK-REBAAKNAKAIIRKFEBAQKAAWAKKAEER 1121
QY 595 KVMASVAVLNKNETCKDYDKFKIPQFLRWFKWGGDDFCERK-----EKIYSFESFKYCECK 650
Db 1122 KKAEEVKAABEERKRIEAEKKAEE-----EERKRIEAKKAEERKRIEAEK 1167
QY 651 KDCDE-----NTCKNKCSEYKWIIDLKKSVEK---QVDKYTKDKNKK--MYNDIDEVKNK 701
Db 1168 KAEERKIIEAAKAEERKRIEAEKKAABEERKRIEAAKAEERKRIEAAKAEERKRIEAAK 1227
QY 702 -EANYVLKESKCEKDVN--FDDKIENESPNVEDMCKCDEIKYLBNEIKYPTKTHDIYD 758
Db 1228 AEAAKAEERKRIEAEKKAABEERKRIEAEKKAABEERKRIEAEKKAABEERKRIEAEK 1282
QY 759 IDTFSDTFDGDGTPISINANINEQSGKDTSNSTGSETSDSPVSEHESPESDAAINVEKLSGD 818
Db 1283 -----KKAEEKKKAAKAAKAEERKRIEAEKKAABEERKRIEAEK 1321
QY 819 ESSSETRGILDINDPSVTNNVNEVHDASNTQGSVSTSDITNGHSSSSSLNRTTNAQD--- 875
Db 1322 ESRRRKKKRLKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1381
QY 876 -IKIGSGNEQSDNQNSHSDNSGSLTIGQVPSDNTQNTYDSQNPRTDTPNALASLP 934
Db 1382 RIEAEKKAABEERKRIEAEKKAABEERKRIEAEKKAABEERKRIEAEKKAABEERKRIEAEK 1440
QY 935 SDDKINEIEGDFSSRDSSENGRGTTSNTHDVR-----RTNIVSERVNSHDFI 982
Db 1441 ABEERKRIEAVKKA-BEERKRIEAEKKAABEERKRIEAEKKAABEERKRIEAEKKAABEERKI 1499
```

QY 983 RGHMANNNAHHOYITQENNGITRGQESAGSNVYKDPKRNPFSS-----ENDHKK-NI 1037

Db 1500 IE--AAKKAEEER-----KAEAVKKGBEVIKGNSNJSETKISNNYETRNDDNSPFKLLDE 1553

QY 1038 QEYNSRDTKGVREBIIKLKSKQNCNNSEYMSYCTYSDERSNSPQSPCSRERKKLCCQISD 1097

Db 1554 EYKSRNIDNRNKIIIGSMSENKCTNDVSSKCYDMKDKISSL-GNCSNDRBKQLCCSISD 1612

QY 1098 YCLKYFNFYSTEYNYCJIKSHKSPYKCFKSEGOSSIPYFAAGGILVVIVLL 1150

Db 1613 YCLNYPDYNKYYDCTKKFSDPLYKCFNSEYSKAVYPAGAGIIMSILIAI 1665

RESULT 13

C71625

variant-specific surface protein 1 homolog PFB0010w - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 09-Jul-2004

C:Accession: C71625

R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.; ; Perlea, M.; Salzborg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O. Science 282, 1126-1132, 1998

A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.

A:Reference number: A71600; MUID:99021743; PMID:9804551

A:Accession: C71625

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1711 <R>

A:Cross-references: UNIPROT:O96108; UNIPARC:UPI0000007BE17; GB:AE001362; NID

A:Experimental source: clone 3D7

C:Genetics:

A:Gene: PFB0010w

Qy	346	-----KKYE-----	349
Db	514	PPENTTKIPKLTAEKRTGILKKYKFCKNSDGNNGGQIKWKECHYEKNDKODGNGDINN	573
Qy	350	-----ACPGGCRLLPC-----	359
Db	574	CIOGDWTKSNVYPIYSYSPFYGSIIIMLNESIEWRBLKSCINDAKUKCKRKGCNPC	633
Qy	360	SSYTSFMKSKTQMEVLTNLYKKKN-----SGVD-----KNNFLNDLFKKNKND	404
Db	634	ECYKRWVEKKKBWDKIKBFRKQKDLLKDIAGWDAGELLEFLYLENIFLEDKMNANGDPK	693
Qy	405	LDDEPKY--EKEVDDLCDBRYTATIIKSPF-----NGPA	436
Db	694	VIEKPREILKENBEVQDPLTKTKTIDDFLEKELNEAKNCVEKNPDNECPKQKAPGDGAA	753
Qy	437	KND-----	439
Db	754	PSDPPREDITHDGEHSSDEDEEBEHEEEOOPAEGTQOEKESKEVVEQETPOKDT	813
Qy	440	-----VDIASOI-----NVNDLRFGCN--YKSNNEKSWNC---TGFTNKPFGT	479
Db	814	EKTVPITPTVDVCTKALADTGSLSNA-ACSLKYVTGKNYGRWCIAISGTTSGKOGAI	872
Qy	480	CEBPRQTLCLGRTYLLHRGHEBDYKEHLLGASIYBAQLLKYYKKEKDNALCSIIQNSY	539
Db	873	CVPPTQELCLYYLKELSDTTQGLREAFIKTAAQETYLLWQYKEDQKNETAS---TE	928
Qy	540	ADLAD--IIKGSDIIKD-----YYGKMEENLKNK-----DKK	572
Db	929	LDIDDPQTQUNGGEIPEDFKRMFTYFGDYRDLFLGRYIGNDLDKYNNNITAVFQNGDHI	988
Qy	573	RNBESLIKIFREKWWENKENVWKMVAVLK--NKETCKD-----YDK	612
Db	989	PNGOKTDROQBEFGYIGKDIWKGMLCALQEAAGKKTLETYNSYNTVFNHGLTGTKLNE	1048
Qy	613	FQKIPOLRWFKWGGDDFCBKREKITYSPESFKVBC-----KKDCDENTCKNKCSE	664
Db	1049	FASRPSFLRWTEWGDQFC---RERITQILKERCMVYQYNGDKGDDKKEKTEACTY	1105
Qy	665	YKMWIDLKSEYKQVDKTKDKNKKMYDNIDVK--NKEANYVLKESKE-----	713
Db	1106	YKMWLTWQDNYKKQRYTEVKGTSFYKEDSDVKESKYAHGYLRKILKNIICTSGTDIA	1165
Qy	714	-CK-----DVNFDDKI--FNEPSNEYEDMC-----KKCEIKYL-NEI	747
Db	1166	YCNMEGTSYDSSNNDNIPESLKYPIEIEBECTCDPSPGVEIPEKKVPEKVLPKPP	1225
Qy	748	KYPKTKHDIYDIDT-----FSDT-----FGDG-----TPISINANINEQQ	782
Db	1226	KLPKROKBERDPTPALKNAMLSSTIMWSIGIGAFATFYLYLKKTKSTIDLLRVINIPK	1285
Qy	783	SGKDTNKGNETSDSPVSHEPSDAAINVEK---LSGDE-----SSSETR	825
Db	1286	SDYDI-----PTKLSPNRYPIYTSYKRYKRYIYLEGSDGSDGYTDHYSDITSSSESE	1339
Qy	826	-GILDIND-----PSVTNNVNEVHDAS--NTQGSVSNT-----SDITNGHSESSL--N	868
Db	1340	YEELINDIYAPAPKYKTLIEVLVLPSPGNNTTASGNTPSDQTQNDIQNDGIPSSKITDN	1399
Qy	869	RTTNAODIKIGRSGNQENSHSSDNGSLTIGQVPSEBNTQ--NTYDSQNPHR---	924
Db	1400	EWNTLKDEFTISQYLQSEQNPNDVNDYSS-----GDIPL--NTQPTLYPDNPEKFP	1449
Qy	925	-----DTPNALASLPDDKINIEFGDSSRDSEN-----	953
Db	1450	ITSIHDRDLYSGEYSYNVNMVNTNNDIPIISGKNGYSGIDILINDSLNSNNVDIYDEVLK	1509
Qy	954	-----GRGDTTNSHDVRRTNIVSERR-----VNSHDPFIRGMANN-----NAHQ	994
Db	1510	RKNELFGTNHHTKNT-----STNSVAKELCGDDPIMNQLDHLHKWDRHDMCKWNKKE	1565
Qy	995	YITQIE-----NNGIIR--GOEESAGN-----SVNYKONPKRSNPFSEN-----	1031

Db 1566 VLDKLEKWKNDNSGNGINSGNINPGNTPTPSDIPSGKSLDTPSDNNIPSSNKTINTD 1625
Qy 1032 -----DHKNIOEYNSRDTKRVREIILKSLK 1057
Db 1626 VSIQHMDNPKPINQFTNMDT--ILEDLEKYNE 1656

RESULT 14
T28626
variant-specific surface protein 2 - malaria parasite (Plasmodium falciparum)
C;Species: plasmodium falciparum
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T28626
R;Su. X.Z.; Heatwole, V.M.; Wertheimer, S.P.; Guinet, F.; Herrfeldt, J.A.; Peterson, D.S.
Cell 82, 89-100, 1995
A;Title: The large diverse gene family var encodes proteins involved in cytoadherence an
A;Reference number: 220487; MUID:95330813; PMID:7606788
A;Accession: T28626
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-2664 <SUX>
A;Cross-references: UNIPROT:Q26033; UNIPARC:UPI000007565C; EMBL:L40609; NID:5886376; PID
C;Genetics:
A;Introns: 2197/3
A;Note: var-2

Query Match 6.7%; Score 433.5; DB 2; Length 2664;
Best Local Similarity 18.5%; Pred. No. 7.2e-11;
Matches 296; Conservative 175; Mismatches 455; Indels 671; Gaps 70;

Qy 16 LYNVIRINESI-----IGRLYNQDESSD--ISRVSPELNHNHNTIY 58
Db 670 LENVLNIDELFDQITEAYGNSQIKQIGKIDTLAKKTKQAADDATEQKNTIDL-----LF 722

Qy 59 DSDVEDVNN-KLINSFVENKSVKKRSLSLFFNNKTKSYDIIPPSYR--NDKFNSEN 115
Db 723 EYDSEEAEEKKIQEBCQPKPTKVRNPGYNN---TYDALAGVAKLOQEAKEQLDRN 779

Qy 116 EDNS---GNTNSNNFANTSEISGDKNKQVTFIQKRLTHLPACGIKRK----- 159
Db 780 DSRSAKANASQGYKSNQGDPPDFKN-----LCGITQKHSNAIGDSKNPCN 826

Qy 160 -----SIKWTCRNSEKITVC-----VPDRKIOLCVANF--LNSRLTEWKFKEIF 203
Db 827 NKGKREFNVGEEKW---KNGGEVKMSHTDLVLPRRQHFTCSNLEHLNTKSTGLTSDKRAIH 883

Qy 204 LISVNTAKLLYNKNEGD-----PSIFCNELRNSFSDFRSFIDGDM-DFGDN- 251
Db 884 SLLGDV---LLAAKKEGEDIKTKLTENDNRSSI CRTWKYSFADIGDIIRGTDLWDINGDA 940

Qy 252 -----TDRVKGYINTKFSDYKKEKNVEKLNINIKKEWEKKNKANLWNHIVN 297
Db 941 TGVQNQLKDIFSKITEELKQHPDKFNDNDKYTNSKHTKLRSDWEANRDQWKAHTCP 1000

Qy 298 HK-GNISKECAIIPAE--POINWIKEWENFLEWKKRFLN-----KDK 341
Db 1001 TKNGNI--QCQATPHDDYIPQRLRWVWEAWEFCQESRLYEELLRDCGSGCTTGKCNNDK 1058

Qy 342 CVENKKEACFGGCLPCSSYTSFMKSKTKQMEVLNLYK-----KNKSGVDKNF 392
Db 1059 CAK-----CDKQCEYKTKIQPWADQWNEISNKYQILYWQAIAINGTEKSTT 1108

Qy 393 LND----- 395

Db 1109 TKDDKDNVDLFLQKLEANYGTRGPPPEAHPDRRPRRAATSKSDVYETTAGYIHOEART 1168

Qy 396 -----LFKQKNQNDLDDFPKKEFYDDLCDCRYT----- 424

Db 1169 RECLGQNVFCNNNGNNEYAFSLTPHEYPKACKCNENKASSPEELGRSDSFDDHQTTPREE 1228

Qy 425 -----ATIIKSLFNGPAKNDV-----DIASQI 446

RESULT 15

H71621

serine/threonine-specific protein kinase (EC 2.7.1.-) PFB0150c - malaria parasite (Plasm
C;Species: Plasmodium falciparum
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 31-Dec-2004
C;Accession: H71621

Db 1229 DEVHSSEGEDESEDEKEEVEEVEHVDGADEKAGAVSQPEASPTTKDVVVKPCDIVKEL 1288
Qy 447 --NVNDLR-----GFGCNYKSNNEKSWNC-----TGT 471
Db 1289 FSNVDLTQACSTKYGPGKNY-----GWRCTPTKTSNDVTGBDGGQSRVVRSTPESGS 1342
Qy 472 FYNKPFCTCEPPRQTLCLGRYYLLHRGHEDY-----KEHLG- 510
Db 1343 NSDKNGATCIPPRRRLLYVGLEQWANKENTSVSQGEATEARGSEAPAPGKESSG 1402

Qy 511 -----ASIVEAQLLKVKYKEDENALCIIQNSYADLAD----- 544
Db 1403 KETPSDKLTARTIESAAVTFFLWDRYKKEWLAKQAEIQNGGLDLYSSGDGDPDPQNK 1462

Qy 545 IING---SDIIK-----DY-----YGKMEENLNKVKDKRNEESLIKPR 582
Db 1463 LLNGVIPPDLRLMFLMFTLGDYRDLVHGNTSDSGNTGNSNNNIVLEASGNEDMQKIQ 1522

Qy 583 E-----KWDENKENVWVMSAVL-----KNKETC----- 607
Db 1523 EKIEQILPKNGGTPLVPKSSAQTPDKWNNEHAESIWKGMICALTYTEKNPDTSGARDENK 1582

Qy 608 --KD---YDKF-----OKI----- 616
Db 1583 IEKDEVEYKFPGSTADKGTASTPTGTYKTQDYKVKLEDTSGAKTSPASSDTPLLS 1642

Qy 617 ----POFLRMFKWGGDDFCERKEIYSPESFVEK-----KK-----KDCDB----- 655
Db 1643 FVLRPFPYFVLEBWGQFCKERRKRL--KQIKECWGSDGDKKYSGDGEQCDRRDTSNEV 1699

Qy 656 -----NTCKNKSEYKKWIDLKSEYKQVDKYTKDKKKQMYDNDIDVKNKENVYLKE 709
Db 1700 SADLEGRSCGNSCRFYKWKIKRKRKYDQANAYSQKTKY-----EGSKGAG--LND 1751

Qy 710 KSKECKDVNFDDKI FNESFNEYEDMCK--CDEIKYLNEIKY PTKTKHDIYDITFSDT 765
Db 1752 HNKE-----FCVLGTCTDAAAFNLRLKNGCPCKD----- 1781

Qy 766 FGDGTPISANINIEQSGKDTNTGNSETSDSPVSH-EPESDAAINVEK---LSGDSS 821
Db 1782 -----NENGNDINFCTEETPRPAENCKPCSSFKINCRNGNCRSGDGD 1826

Qy 822 SETRGILDINDPSVTNNVNEVHDASNTQGSVNTSDITNGHSESSLNRTTNAODIKGRS 881
Db 1827 KE-----KCGGTITTG-----NFTMTGTC--TEDVWVHVSVDKANEFEG----- 1864

Qy 882 GNEQSDNQNSSSHSSDNGSLTIGQVPSEDNTQNTYDSQNPHRDTNALASLPSDDKINE 941
Db 1865 -----DGLDEACENAGIFT--GIRKDEWKCKVCGLHICKOEKNG-----AIND 1907

Qy 942 -----IEGFDSSRDS-----ENGRGDTTSTNTHVRRRTNIVSE----- 973
Db 1908 QQIILVRALLKRWVEYFLEDYKKIKKLKPCIEENGSGTCTINGCN-KKCNRVGEWIKLKK 1966

Qy 974 ---RRVNSHDFRN--GMAN-----NNAHHQYITQIENNGIIRGO-----ESAGNSVN 1017
Db 1967 DEWTKIKNHYLEKNKEGDKNVTSLVTNVLFTLVTQIAAANDKREQTSLDLKTSLG--CN 2024

Qy 1018 YKDNPKRSNFSSND-----HKKNIQEYNSRDTK-----RVREI 1053
Db 2025 CPEN-SRKDNGNDNDAIDCMLNKLETKIHECKTQHENSSENSQPHNCGNPPPPDEDL 2083

Qy 1054 KLSKQNKCNNEYSMBEYCTYSDERNSSPGPCSRERKK 1090
Db 2084 LLEENPEVQP---GFCPTPQQPPEPDKGKLEKK 2117

Db 752 KAGGTGASGTGPKNIMDKLIEHELQEAKKCKDCQEPQOSLGRSLNPHVVDGSPKKR 811
Qy 423 -----YTAT-----IIKSFNGPAKNDVD-----441
Db 812 DKRTNPCYSDTTTEYAVLAGVAKQFQGEVRAKMLERSRKNGETKSSLEGDIKKAQFNG 871
Qy 442 -IASQIN-----VNDLGF-----GCNYKSNNEKSWNCTGTFINKPGTC 480
Db 872 RSGSELNGDICKINDKYNDSINSTAGGPTCGDKGNERFNAGTKWEGDNFVATHKNLY 931
Qy 481 EPRRQTLCL--LRTYLLHRGHEEDYKEHLG---ASIVEAQLLK--YKYKEKBNAL 531
Db 932 IPPRQHMCTSNLEKDLFVTSKSNVNDPFLGDLVLAANNEAORTKXFAHKDODHGA 991
Qy 532 CSIIQNSYADLADIIKGSDI-1KDYQYKGMKEENLKV-----NKDKRNE 575
Db 992 CRSVRYSPADLADIIRGRDMWDKDGAKMEDIPKIFGNLYESLPGIKGYDGDQRT 1051
Qy 576 ESUKIPREKWDENKENVKMSAV---LKNKETCKDYDFQKIPQLRWFKEWGDGDFC- 631
Db 1052 Q-YKQREDWNEARDQVWAMWCEKDGIIKDEDPVDDY--IPQRLRWMTWAEWYCK 1108
Qy 632 ---EKREKIYSPFSFVKE-----CKKDCDENTCKNKSEYKMWIDLKSEYKOVDKY 683
Db 1109 VQSEYDELLKCKGCKIKGVQCTSGDSCTPCAEACTTYGKIKPWSQDQWNMLQY 1168
Qy 684 T-----KDNKMYDINIDEVKNKE-----ANVYLKEKSK- 712
Db 1169 TLLYQWAEATTARYGGTRAYSGVDGDKPVVQVLEELQKQNSGKTTVNTAAGYIHEARV 1228
Qy 713 -ECK-----DVNPDKIPNESPNEDYCKKCKDEIKY-----LNKIKPKTKHDI 756
Db 1229 GECEVQYFCNTNGNQDKYFREPDKDHEACKDRPOOSAGGAGARSLPSR-----1283
Qy 757 YDIDTFDITFDGTFPISINAMINEQOQSGKOTSNTGNSSETSDSPVSHPEPESDAAINVEKLS 816
Db 1284 -PVDSDDH-----SSEDEDEEDEDGDEGAEDENDPEASEVKDDEDVBEETA 1333
Qy 817 GDESSSTRGILDINDPSVTNNVNEVHD-----844
Db 1334 VSQPAAPT-----TTTGVTPACBIVKDLPEKPKNTFKEACTQYGGNNSRLGWKCIPTS 1388
Qy 845 -----ASNTQGSVENTDITNGHSESLNR-----TTNAQDIKGRSGNOSDN-----888
Db 1389 GGDKAATRGSGDITTKQNDSEGESEGHQRAKRTSDASGEKSAKSAGTGGTICIPPR 1448
Qy 889 -----QNSHSSDSSSLTIGQVPSDNTQNTYDSONP-----922
Db 1449 RKLVGGLTKWAEIQSSQSALSGQTTAGTPSQ-----AQDPLLAFAVESAAVETP 1500
Qy 923 -----HRDTPNA-----LASLPSD-----936
Db 1501 FLMDRYKLNAPQSGSLGGAPQLLNGAIGSEBETPETSLSKGTIPPDFLRLMFTYLG 1560
Qy 937 -----DKINEIEGFPSSRSENGRGTTSNTHDVR-----TNIVSERV 976
Db 1561 YRDICVGKQDVIKALEASGDKSSKNPQOEISSIKIEILKNGGTPTTPVTHSPSGTT 1620
Qy 977 NSHDFIRNGMANNAHQYITQIENNGII-----RQGESAGNSVNYKDN-----PKRSNF 1027
Db 1621 PSSHWTNG-----QHIWK-----GMICALTYESGEKKIEQVKATDNTDLFEKLKDK 1668
Qy 1028 SSENDHKKNIOEYNSRDTKRVRE-----EIIKLSKQNKCNBYSMEYC-----1070
Db 1669 YSDYD-KVEIKEENETEAKGPDQGLTPTOTLLSNFVRKPPVFRYLEEWGQNFCKTRKML 1727
Qy 1071 ---TYSERNSSPGCSREERKLLCCQISDYC-----LKYFN-FYSIEYCNICKS 1116
Db 1728 KDIIECRNSDNP-----HDYCSGDGYCHCNELKHNMFHIDCRDCKY- 1772
Qy 1117 EIKSPYKCFKSEQSSIPVFAAGILVIVILLSSASRMKSNBEYDIGES-----1168
Db 1773 -----ECRKYKWHKXDFEH-----NQNKYGEEHEKLTNGDNGYGGGDNT 1815

Qy 1169 ----NTEATEFENNYLKLRSIFNQEVQETNISDYSEYN 1203
Db 1816 NFCQIKKKKTAEDFLKALRHCKNSE-DDTDKSEDEKEN 1853

RESULT 19

T18477

hypothetical protein C0485w - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T18477

R:Lawson, D.; Bowman, S.; Barrell, B.

submitted to the EMBL Data Library, November 1998

A:Reference number: Z18937

A:Accession: T18477

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-2523 <LAW>

A:Cross-references: UNIPROT:O77365; UNIPARC:UPI000017CC31; EMBL:AL008970; NID:e1407852;

C:Genetics:

A:Map position: 3

A:Introns: 148/3

A:Note: C0485w

Query Match 6.1%; Score 395; DB 2; Length 2523;

Best Local Similarity 19.7%; Pred. No. 3e-09;

Matches 251; Conservative 211; Mismatches 458; Indels 354; Gaps 56;

Qy 32 YNRQD-----ESSDISRVNSPELNNHKTNI-YDSYEDVNNK-LINSFVENKSVKKRSL 85

Db 1359 YNKKEYMIKKGESSNVPHKHKNNKKNYCYNDLGMHSLQNRHTITSEVSSKFLCKNMK- 1417

Qy 86 SFNNKTKSYDI--IPPSYSYRND---KFNLSLSESDNSGNTNSNNTNFANTSEISIGKDNK 140

Db 1418 NPYDKSNNEIEIHKISASNIPIRHTMVCVSNKGNNGNNINNYKGPATKALV-----NK 1472

Qy 141 QYTFIQKRTLHPACGIRKRSIKWICRESEKIVTC---VPDRKIQLCVANFNLNSLETW 196

Db 1473 --LFISKKE-----SKRAITSSKKRDDNINVIKINTPQSKVSEKRNNNNNNVLG 1523

Qy 197 EKFK-----EIFL--ISVNTAEKLLYNKNEG--KDSIFCNELNRSFSDFRSFIQDDMD 247

Db 1524 DKNKNNDDELFTKEIKKSTISKQKKGNEGTTHKDNINILNEDVDHFKQPSLREVT 1583

Qy 248 FGGTDRVKGYINTKFSYVYKEKVEKLNKKEKWEKKA-----288

Db 1584 KKNKKNKKNKKNKKNKKNKKNKKNKKNKKNKKNKKNKKNKKNKKNKKNKKNKKNKKNKKN 1643

Qy 289 -----NLNWHMIVNHKGNISKECAIIPABEPPQINLWIKENWNFLMEKKGLFLNKKCV 343

Db 1644 CYKNGDIYNMIEIINVDNIHK-----IDKKGNDIDYKDKSL 1679

Qy 344 ENKYEACGGCLPCSSYTSFMKSKTQWE-----374

Db 1680 ENNKQMKGLIKILPSSYTEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKE 1739

Qy 375 -----VLNLYKKNSGVNDKKNFLND-----LFFKKNKN 403

Db 1740 VKNKLHLKRETCVTCVNDVNIQNKKGKDDKKNIKRDQHVGLKFLDEMSAMFEKKKKI 1799

Qy 404 DLDDPFNE--KEYDCLDCRYTATIIKGLFNGPAKNDV-----DIASQ 445

Db 1800 KKDDINKKEDINKKDDINK-----KKNIN--KKNINKKDDINKKDDINKKDDINKK 1849

Qy 446 INVNDLRGGCNVYKSN-EKSNWCTGTFTNPGTCEPPRQTLCLGRTYLLHRGHEEDY 504

Db 1850 DDINKKNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1900

Qy 505 KEHLGASIEAQLLYKYKEKEDENALCSTIQNSYADLADIIKGSIIIKDYGKMEENL 564

Db 1901 -----IGLKKKKIKRKNVTTTIATKHDNIIDI-----KKKNEKE 1934


```

770 QY TPIINANINEQQSGKOTSNTGNSSETS-----DSPVSHE--PESDAAINVEKLSDGES 820
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
493 Db SGGTSGA-----SGTDENKGTFYRSEVQCPCDGVQHKGGNWERKTVKCKMRWSKL 546
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 821 SSETRGILDIINDPSVTNNNVNEVD-----ASNTQGSVSTNDITNGHSSSS 866
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 547 YKPINGKMVL-----LLSKLVKKDMMLKNWKEFCLTONSSDGSVG--SVVTTGASGN 600
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 867 LNRITNAQDIKIGRSCNQSNQSHSSDN----SGSLTIGQVPSEDNTQTNTYDSNP 922
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 601 SEKKELVDWKCYKHNEVQKVNOGEVEDDELKGAGGLCILPNPKKNKEVSEAKSQNN 660
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 923 HRDTENA-----LASLPSDDKINEIEGFDSRSDSENCRGDTT 959
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 661 HADIQTPHPFFYYVAHMLKDSIHWRTFKLSKCSISDKTMK-----CRNCNKKCDCFP 714
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 960 SNTHDVRRTNIVSERRRVNSHDPIRGMANNNAHQYITQIENNNGIIRQG-----EESAG 1013
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 715 EKWVKQKET---EWKPIKDHFKTQSGIPEG---YYFTTLEL-ILKLQFLKEDTEENTE 765
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1014 NSVNKYKONPKRSNFS--ENDHKNIQBYNS-RDTKRVREBII--KLSQNKCNKEYSME 1068
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 766 NLSDAEEABELKHLOKLIKENENNLAVNAGTEOKLMDKLLNELNDATKCD-----820
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1069 YCTYSDE-----RNSSPGP---CSRERKK 1090
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 821 -CPIDPEEDKSGRSADPSDFIPRPERKE 849
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 24
T28317
ORF MSV156 hypothetical protein - Melanoplus sanguinipes entomopoxvirus
C:Species: Melanoplus sanguinipes entomopoxvirus
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C:Accession: T28317
J. Virol. 73, 533-552, 1999
R:Afonso, C.L.; Tulman, E.R.; Lu, Z.; Oma, E.; Kutish, G.F.; Rock, D.L.
A:Title: The genome of Melanoplus sanguinipes entomopoxvirus.
A:Reference number: Z20484; MUID:99102612; PMID:9847359
A:Accession: T28317
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1127 <AFO>
A:Cross-references: UNIPROT:Q9YVT6; UNIPARC:UPI00000F6900; EMBL:AF063866; NID:g4049647;
C:Genetics:
A>Note: MSV156

Query Match          5.7%; Score 366.5; DB 2; Length 1127;
Best Local Similarity 21.8%; Pred. No. 1.9e-08;
Matches 293; Conservative 176; Mismatches 417; Indels 456; Gaps 66;

QY 50 NNNHKTNIYDSGYEDVNN-----KLINSFVE--NKSXVKKGSLSFINNKTASYD---IIP 99
    ||| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 25 NNKVSLDIINSLYETLNINKIFSDKITNEIKYNKIVEK--IFYMHOFKINDYNIILQ 81
    ||| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 100 PSYSYRNDKFNSLSENEEDNSGNTSNSNNFANTSEISIGKDNKOYTFIQKETHLPACGIKRK 159
    ||| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 82 YLIEYNNINKCIKENCFCKRNPLYN-----ITY-KKKLYIYDLDYEEK 124
    ||| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 160 SIKWTICRENSEKITV-CYPDRKIQLCVANFLMSRLTMEKFEKIFLISVNTAEKLLYNKN 218
    ||| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 125 KDKELVINIEQNADVKNIDKNVY--NNIHSNDETITGETIIDLINKLLKLVSDE 181
    ||| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 219 EGKDPSIFCNELRNSFSDFRSFGIDMDFGNTDRVKGYINTKPSDYKYEKNVPLNNI 278
    ||| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 182 KQIEQIYKYN-INNKIEBFK-----NIDNVQKEINKQDELNKLDESKEEFI 228
    ||| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 279 KKEWEKKNKA-----NLWNHNVHNKGNISKECALIPAEEPQINLIWKEWNE----325
    ||| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 229 KKQ-EELNKTIDKQEELIKLNDKEINP-NIDBKQLL-----DOJNSKINTLNTENK 281
    ||| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 326 --NFLMEKRIFLNTRDKCVENKTYEACFGGRLPCSSYTFPMKKSKTOMEV-----LT 377
    ||| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```


Db 323 -----TFVI---NNKFLISENNILLEDKISFIDRKIES--NKCEDYCVANNNNN 368
Qy 368 KSKTQMEVLTLNLYKKK-----NSGVDKNNFLNDLPK----- 398
Db 369 BRNLSLILENAYKDCESRTINEDRIYNNFEDMDKISHDAPDFIIPSGFNKEBDNGNEK 428
Qy 399 -----KNNKNDLDDP-----FKNEKE-YDDLCDCRYTATIKSP--LNGPAKNDVD 441
Db 429 YQNVFDSKNLENINVEDPPFSFSEKQFFQNCDSNENLWLNKKFDEHNVEKNEIY 488
Qy 442 IASQINVNDLRGFCNYK-----SNV-----EKSNNCTGTFNNKPPGCEPPRR 485
Db 489 EPKNVYENE-----NYDQKDVBSSKFFENNVPFDDKNKNVDRIIDSGVEGNCVBEK 542
Qy 486 QTLCLGRYLLHRG-----HEBDYKE-----HLIGASIVYEAQLLYKKYK 524
Db 543 LKKEBEKTYFVETGNYGDELPRRNFEEIDENYKEVBEKFDKMGNGENFFEE--VEEKYD 600
Qy 525 EKDNALCSIIQNSYADLADIIGSDIILKYGKQMBEN---LKNVNDKKNNEESLKI 580
Db 601 EK-----VGKNIFE-----VEBEKDEKRENIPIABIEEYDEKMGNIPEE 643
Qy 581 PREKWDENKNNVWKMVAVLNKK-----ETCKDYDKFKIPQFLRFKMGWDDPCCK 633
Db 644 VEENFGEKVGKNIPEEVEBEKFDKMGEPFFDEVEBEKFD--EKWGEF--FFDEVEBEKPBK 699
Qy 634 RKEKI-----YSFESFKVECK-----KDCDENTCKNCSYKKW 668
Db 700 MGENIFESIPKODVEIBETYSERKMWETHDKIKDKYDEPEYKIHDE---KKEVEEFPFI 756
Qy 669 IDLKKSEYE-KQVDKYTKDKNNKMDN-----IDB---VKNEANVYLKESKECKQVNF 719
Db 757 ADKKEENEDSNVELINDKNFPENKETEIDEKVSKNMBEEDPVYENNETFECEDIFL 816
Qy 720 DDKIPNESPNYEDMKCKDBI-----KYLNKIKYPTK 753
Db 817 KREDNDSENE-----KEIDBIEGINIKYHLNKNNSYDDVHLTHDFKNELLIEKN 871
Qy 754 HD-----IYDIDTF--SDIFGDTPIISIANI-----NEQSGKDTNTGNET 795
Db 872 VDNICSDDNIYDGNICGDDNIYDGNIDYDGNIDYDGNIDYDGNIDYDGNIDYDGNIDY 931
Qy 796 SDSPVSHPEPDAAT-----NVEKL-----SGDESSSETRGILDNDPST 836
Db 932 GDNISGNIIDDDNIYDGNIDYDGNIDYDGNIDYDGNIDYDGNIDYDGNIDYDGNIDY 991
Qy 837 NNVAEVDHAS-NTQGSVNTSDITNGHSESLNRTTNAQDIKIGSGNEQSDNQNSHS 895
Db 992 DYVEENSDFVYDKGGMWVNELIGEYSEKYMNNIEDNELVIW-SASVKDKERLNDN 1050
Qy 896 SDNGSLTIGQVP-SEDNTQTYDSQNPRTPNALASLPDDKINIEGDFDSDSDSNG 954
Db 1051 IDLANNISNDYIKNNEDIKNVHDS-----FSISNKSSELHDLINGILEKSISSN- 1098
Qy 955 RGDITS-----NTHDVRTNIVSERRVNSHD-----FTNGMANNNAHOYITQ 998
Db 1099 --DIKSIEVCVCKENEIHHKNNMKKKSLNNDNNLNDENMYCDISNDIFKNNEYTKHDD 1156
Qy 999 I-----ENNGIIRGOEBSAGSNVYKDNPKRSPFSSNDHKQNTQYNSRDTKRVBEI 1052
Db 1157 VYTPDENNSNLIIGEDHECHVSSNMF-----YFNIISKNTESNNIL-YEQNDKKTNS 1212
Qy 1053 IKLSKQNKNEYSMEYCTYSDERNSSPGPCSRBEERKKLCCQISDYCLKYNFY--SLEY 1110
Db 1213 VK-----HPMTY-----IKGFEYASDSINF 1232
Qy 1111 YNCIKSEIKSEYKCFKSEGOSSIPYFAAGGLVVLVILLSSASRWGKSNEEYDIGESNI 1170
Db 1233 LKALKGLPPLFLKC-----KDMKPYMLFNLVLKVI-----ESNDYN-GRKRI 1275
Qy 1171 EAT 1173
Db 1276 KVT 1278

RESULT 28

T18402

asparagine/aspartate rich protein - malaria parasite (Plasmodium falciparum) (fragments)
C:Species: Plasmodium falciparum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18402
R:Barale, J.C.; Candelle, D.; Attal-Bonnefoy, G.; Dehoux, P.; Bonnefoy, S.; Ridley, R.;
Infect. Immun. 65, 3003-3010, 1997
A:Title: Plasmodium falciparum AAP1, a giant protein containing repeated motifs rich in
A:Reference number: T18929; MUID:97378065; PMID:9234746
A:Accession: T18402
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-3844 <BAR>
A:Cross-references: UNIPROT:Q94648; UNIPARC:UPI000007ABAB; EMBL:Y08926; NID:e1154302; P1;
C:Genetics:
A:Gene: aarp1

Query Match 5.5%; Score 357; DB 2; Length 3844;

Best Local Similarity 19.0%; Pred. No. 2e-07;
Matches 302; Conservative 226; Mismatches 552; Indels 510; Gaps 70;

Qy 9 FLIPLIPLYNVIRINESIIGRTLYNQDESSDISRVNSPELNNNHKTN---IYDSDYEDV 65
Db 309 FHVMIIRKMCVLIKIFHVVIILKILNK-----MYNNKNNKELTYNDEY--- 351
Qy 66 NNKLINSFVENKSVKKSRSISFINNKTYSYDIIPPSYSYRNDKFNLSSENEDSGNT--- 122
Db 352 ---IINEFIRDFYSIEKHSBEIINOR-----KILQERYENDGNMKI 390
Qy 123 -----NSNNFA---NTSEISIGKDNKQYTFIQKRTHLFACGKRKSI 161
Db 391 TLFHFKKFPTSDDIYLESNYFKWIMNNSSKRECNKRYIKYRET-----IQEKNI 444
Qy 162 KWI CRE-----NSKTIWCVDPDRKIQLCVAN---FLNSRLRTM 196
Db 445 KIKCNIIYIEKLNKSYEFNEFLDIKYDKEKHNIIHLPLIYDLYMGDIFEFNL--LPDL 502
Qy 197 EKFKEIFLISVNT-----EAKLLYNKNEGKDPISFCNELRNSFSDFRSPFGDD---M 246
Db 503 KLLKNILPINIQLTFYIEIKYGVWRNGPQMIQVGEYDNSLF-FQNDLIGIQFALLMM 561
Qy 247 DFGNTDRVKYINTKSDYKE---KNVEKLANIKKEWEKKNKAN-LWNEHMIVNHKNI 302
Db 562 NVLPYENRKVYLDHSHVINLFHQIWNIKVQYDEFLKXLPMSKNYNDVQNDYI--HKAD- 618
Qy 303 SKECAIIPABEPQINLWIK-----EWSEN---FLME-----KKRLFLNIDKCKVENK 348
Db 619 DEICINEKYDVEDTNKYIDTSPNFSYNNHINQNHVWVEHSEDKKPYNN-KIKYI--KKN 675
Qy 349 EACFGGCRLPCCSYTSPM-----KSKTQMEVLN-----LYKK 382
Db 676 DEFP--BEHMKYEMSLIYNRSKSDNNNTLPVNNKMGNSKGNLDDDDGHVKLLFGI 732
Qy 383 KNSGVKQKNNFLDLFKQKNL-----DDFKNKEKYEDDLCD--CRYTATIKSFLN- 433
Db 733 DDKNVNNNNNTNINSINNINNVNNNNNTYVYKVEFKSICDLCKYSCKQLKQDENE 792
Qy 434 -GPAKNDVD---IASQINVNDLRGFCNYKS---NNEKSNWCTGTFNKF---PGTCEP 482
Db 793 SNPLRTEGRLTSDNINFNKLDNNDNIQDLDNDKRYNKGVSVRKESYNEEDRICEK 852
Qy 483 PRRTCLGRITLLHRGHEEDYKEHLIGASIVYEAQLLYK--YKEKDNALCSIIQNSYA 540
Db 853 YKSKSL-----EWNEYLEEFWNIEIRNPFILYQILFKNFRMLNIRKINDMYV 901
Qy 541 DL-ADIIKSGD-IKDYVYKQMBENLKV-----NKDKK 572
Db 902 EIEIHHKMAKDFSDIYTNKSTVDSSKIGHIFQFFYDIIIRIFNELYLYESNIPRK 961
Qy 573 RNBSLSKIFREKWDENKENVKMSAVLKNKCKDYDKFQ-----KIPQFLRW 622

E71609	Ser/Thr protein kinase PFB0665w - malaria parasite (Plasmodium falciparum)	
C;Species:	Plasmodium falciparum	
C;Date:	13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 09-Jul-2004	
C;Accession:	E71609	
R;Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;		
; Perlea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.		
Science 282, 1126-1132, 1998		
A;Title:	Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.	
A;Reference number:	A71600; MUID:99021743; PMID:9804551	
A;Accession:	E71609	
A;Status:	preliminary; nucleic acid sequence not shown; translation not shown	
A;Molecule type:	DNA	
A;Residues:	1-1714 <GAR>	
A;Cross-references:	UNIPROT:O96226; UNIPARC:UPI0000082380; GB:AB001409; GB:AB001362; NID	
A;Experimental source:	clone 3D7	
C;Genetics:		
A;Gene:	PFB0665w	
Query Match	5.2%; Score 334.5; DB 2; Length 1714;	
Best Local Similarity	19.1%; Pred. No. 7.3e-07;	
Matches	306; Conservative 204; Mismatches 551; Indels 537; Gaps 73;	
Qy	5 FN-IYFLI-----PLIFLYNVIRINESI-IGTLYNQDESSDISRVNSPE-----48	
Db	31 FNDIYEIITNHNKQPHIKENNIKIYTRNVNDRSLVDEKXKNDINNIDKYEKTTCSY 90	
Qy	49 -LNNHKTNIYDSYEDVNNKLINSF-----VENKSVKGRSLSPINNKTYSYDIIPP 100	
Db	91 VLNNLHK-----KYNHNNKMYDEYFYDEYLINIKLKGKFNVIIEGRKGNRLG 144	
Qy	101 SYSVRNDFNLSNENSGNTNNGNNTFANTSEISIGKNDKQYTFIQKTHLFAIGIRKS 160	
Db	145 VSSTSDNK---KKNKKRYNNNNN-DNNNDINNDCNNKNY-----PCCS-----186	
Qy	161 IKWLREN---SEKITVCVPDRKIQL--CVANFL-----NSRLETMEKPEI-----202	
Db	187 ---SCGNVLSSSTFNWCEGDKKISYGRQITNLVSCYKNQKLSYNTIHTINQQVHDN 243	
Qy	203 -----FLISVNTAEKLLY-KNNEGKDPISFCNELNSFDRSSFIGDD-MDFGGNTD 253	
Db	244 NIYVDNQHMLYNTDNLKSYNNKMDLSYNLHEKNSFSNFINSFGRNPHELCCKLK 303	
Qy	254 RVKGYINTKFSDYKERN-----VEKLN-----NTKK-----EWEKRNKANLWN 292	
Db	304 KAVEY-KERVVIDINKEDFVLLGISKTCVKKCNKCSGDVNTKIDKCVDEEKSKEG---359	
Qy	293 HMIYNHKGNTSKCAIIPAEEPQINLWKENENFLMEKRLFLNI-----KDKC 342	
Db	360 -VILNY-----MKDILFYNTFRNNDNPNRKEKP 388	
Qy	343 VENKYEACFGGCRLPSC-SYTFPMKSKTOMEVLTNLYKKNSGVKNNFLND-----395	
Db	389 KECQKYNK--DDVHVLCDDHFSKSHSTTKNSNTKLYNVKEXHIHINKYNNVYFVEG 446	
Qy	396 ---LFKNNKNDLDFPKEKEYDD---LCDCRYTATITKSLFNGPAKNDVDIASQ---445	
Db	447 QEKLYSPSIKETQFYQNDYKDDNVKMLSYNYNDWVYKN-----SKGMIDSLSTQHA 501	
Qy	446 -----INVNDLRG-----FCGNY-----458	
Db	502 PKGETVINLKLRRRTSIMRKVYSVLVYFAPGAPWLKIRRGQIKGQKGGKKKNDEN 561	
Qy	459 ----KSNNEKSWNCTGFTNKPFGT---CEPPRQTLCLGRYTLHLRHGHEDEYKEHLIGA 511	
Db	562 KKKNNKNNKNNNNNNINNHGKRVIQYTDKEIQNDYCKNKESSKRGHKKWMEKKNLS 621	
Qy	512 SI-----YEAQLLKVYKEKDN-----ALCSTIIONSYADLAD-----II 546	
Db	622 SLLSINGKCYNKWKNTKTRKPNKGRKGEKYTCYENIKILEDDIKDRFPNDHKNRNL 681	
Qy	547 KGSDIILKDY-YGKQWEENLNKWKDK-----KRNESLKIPIREK 584	

Db	682 NEENFTKEHQINGRNEKHEVNEKNKEEDTFNISKENTKEGSIYIITHKNKRNMDNIKGRYD 741	
Qy	585 WMDENKENTWKMVSAYLVKNKETCKDYDFQK--IPOFLRWFKWEGDDFCCKRKEKIYSFE 642	
Db	742 NINDKKEPSSNLYKCVKN-----DKINKSQTSLFFEFMKGKG-----QKHNVIKKE 790	
Qy	643 SFKVECKKDCDENTCNKCKSEYK-----KWIDLKSEYKQVDKYTKDKNKKMYDN 694	
Db	791 DVEIKFRTNKSPTLTKKISDYKCNLLYTSLDRIHKVNSIYNERIER-TKHVPQKKNND 849	
Qy	695 ID-----EVKNKEANVYLKESKECKVDNE--DKLINESPNEYDMC--735	
Db	850 IDIRGYKSYNFPKSNMNMNSUKCY---HTKTCDSYNDYFMKNKMSKKAQNLVSKCIS 906	
Qy	736 ---KKCDEIKYKNEIKYPTKHDIDYDFSDTFPGDTPISINANTINEQSGKDTSTGN 792	
Db	907 KYKKKAIKKKERKEITTTKKY-IYRKNELIS-----ISPQGVFGHENRRTKE--N 955	
Qy	793 SETSDSPVSHEPESDAAINVEKLSGDESS-----ETRGILDINDPSVTNNVNE- 841	
Db	956 NKSKEAYTSKSRKN-----NKIKGEKTKRSLCSYKLRKMKHLCVENKWHIKKNVRQI 1010	
Qy	842 -----VHDASNTQGSVNTSDITN--GHSSESLARTTNAQDIKIGRSCNEQSDQENS 892	
Db	1011 IKKKKNIIYTKIKLNSYKTLIDQVNVKGDEEHKLSNHNKVKKKKKKNCINEKNDNNND 1070	
Qy	893 SHSDNSGSLITIGQVPSSEDTQNTQSDQPHRTPNALASLPDSDDKINEIEGDFSRSDE 952	
Db	1071 NYNDNDDNNNDNY--NDNNNDYNDNNNDNDHN-----NDNNNDL-----NN 1113	
Qy	953 NRGDTTSTNTHDVRRTNI-----VSERRVNSHDFIRNGM---ANNAHHQVITQ 998	
Db	1114 DHRDNNQREHSCEEETNIQNVQKCEGEKYEGRKNKYTYNNYKINGKNEIHDDY---1170	
Qy	999 IENNGIIRGOEESAGNSVNY-----KDNPKRSN-----PSSEND-1032	
Db	1171 -----NIKSHGSRINTYNIPNIKDNKHNNDKGEKSCELKCKSIPYVKEKYNLENT 1221	
Qy	1033 -----HKKNIOEYNSRDTKRVREIILKLSKNK-----1060	
Db	1222 YEIIGLIYYGDKSQVYKICINNNKRVYAMKVLKECNEIFVDNFIKKYFLKNKPNKNI 1281	
Qy	1061 -----CNNEY-----SMEYC---TYSDERNS-SFGPCSRBERKKLCCQISDYCLKYNFYS 1107	
Db	1282 SIYDIFCNWYIICIIDYCEGSTLLDYFMSLVPGSLDVVEIKKMKNI-----FIA 1332	
Qy	1108 IEYV---NCTKSEIK-----SPEYKCFKSEGOSSSIPYPAAGILVVIV 1147	
Db	1333 LDFFHNNIIRDIKLEINFMKKNKKRKFNYEKGSGFLNNHHEISPTSCSNLHKD 1392	
Qy	1148 LLLSSASRMGKS-----NEEY---DIGESNIEATFEEN-----1178	
Db	1393 LQLRGMDTIGKKMGKKFIRNLYNEKKHNLNIFQKNCSHILLKKNLKNILSNDIOLKS 1452	
Qy	1179 -----YLNKLSRIFNQEOVQETNISDYSEYVNEKMY 1210	
Db	1453 PKCYIKYNNMDTLFNYE-----DGSNWSYNSICY 1483	
RESULT 35		
H71609		
hypothetical protein PFB0630c - malaria parasite (Plasmodium falciparum)		
C;Species: Plasmodium falciparum		
C;Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 09-Jul-2004		
C;Accession: H71609		
R;Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;		
; Perlea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.		
Science 282, 1126-1132, 1998		
A;Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.		
A;Reference number: A71600; MUID:99021743; PMID:9804551		
A;Accession: H71609		
A;Status: preliminary; nucleic acid sequence not shown; translation not shown		
A;Molecule type: DNA		

A:Residues: 1-1166 <GR>
A:Cross-references: UNIPROT:O96219; UNIPARC:UPI000007D874; GB:AE001407; GB:AE001362; NID
A:Experimental source: clone 3D7
C:Genetics:
A:Gene: PFB0630C

[illegible]

```
Qy      879 GRSG--NEQSDNQENSHSNDGSLTIGVPSDNT-----QNTYDSQNPHRDT 927  
         :|:||||:::||::||::||::||::||::||::||::||:  
Db      839 GENGTYDENSTYDCNCTY--DKNRITYDENRTYDNKRTYDENRTYDYDENRTYYDD-- 894  
         |||||::||::||||::||::||::||::||::||::||:  
Qy      928 NALASLPSSDKINEIGFSSRDSENGRGDTTSHTHDVRTNVSVRRVN-----S 978  
         |||||::||::||::||::||::||::||::||::||:  
Db      895 KSCVFHKODIIINIE-EECEKTKEADH---RVNETDDIQLQALLKKKKIREEYTQTPK 950  
         |||||::||::||::||::||::||::||::||::||:  
Qy      979 HDFIRGMANNHAHQVIITQIENNGLIARGOEAGSNVSNY----KNPKRSNPSSENDHK 1034  
         |||||::||::||::||::||::||::||::||::||:  
Db      951 SDISINNKLQDNCKHEY----ENFNHL-DDESTYDDLVDHFTHDDELENONCFSNSNNVK 1005  
         |||||::||::||::||::||::||::||::||::||:  
Qy     1035 KNTOEY-NSRDTKYRBHEI-IKLSKOKNCNMYSMEYCITYSDERN-S-SPGPCSEREKK 1090  
         |||||::||::||::||::||::||::||::||::||:  
Db     1006 MNERNKYIYGRRNLAVENLSQLYVLWDIFTNNIS-KYTVHHFEKNEFIVPKAINEESERK 1064  
         |||||::||::||::||::||::||::||::||::||:  
Qy     1091 LCCQISDYCLKUYNFYSISB-----YNCKIKSEI-----KSPEYKCFKGE 1129  
         |||||::||::||::||::||::||::||::||::||:  
Db     1065 -----RNFFIVNSQMPIYINCISIIVNICRFELFKPLIPKCVIK----- 1109  
         |||||::||::||::||::||::||::||::||::||:  
Qy     1130 QOSSIPPYPAAGGILVVIVL-----LIASSARMG-KNSEEDYDGIESNIATFEENLYL 1180  
         |||||::||::||::||::||::||::||::||::||:  
Db     1110 -----SIICVTAIMAKLHKPHLIPSSEMINKAEDYLIIENKIDQ--EG--L 1153  
         |||||::||::||::||::||::||::||::||::||:  
Qy     1181 NKLSRIFPNQ 1189  
         |||||::||::||  
Db     1154 NELCFLPFQ 1162  
         |||||::||::||
```

RESULT 36

T28677
rhostry protein - Plasmodium yoelii
C:Species: Plasmodium yoelii
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T28677; C45521
P:Keen, J.; Sinha, K.; Brown, K.; Holder, A.
Mol. Biochem. Parasitol. 65, 171-177, 1994
A>Title: A gene coding for a high molecular mass rhostry protein of Plasmodium y
A:Reference number: Z20508; MID: Y9021522; PMID: 7935623
A:Accession: T28677
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-2269 <KE>
A:Cross-references: UNIPROT:Q26223; UNIPARC:UPI0000080871; EMBL:L27838; MID:g45<
P:Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.
Mol. Biochem. Parasitol. 42, 241-246, 1990
A>Title: Identification of the gene for a plasmodium yoelii rhostry protein. Mu
A:Reference number: A45521; MID: 91101660; PMID: 2270106
A:Accession: C45521
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 2131-2269 <KE2>
A:Cross-references: UNIPARC:UPI000017B646; GB:M34283

Query Match	5.1%;	Score 328.5;	DB 2;	Length 2269;
Best Local Similarity	21.8%;	Pred. No. 1.8e-06;		
Matches 315;	Conservative 203;	Mismatches 471;	Indels 457;	Gaps 83;
Qy	16	LYNVIR-----INESIIGRTLYN-----RQESSDISRVSPELNNHKTNIYDSYE- 63		
Db	167	IYNTIKSYFQIYEGDID-TFYNELSSIVKEDPIDDIETKTLELNRSKIDNVDYKIQKM 225		
Qy	64	-----DVNNKLNISFVZ-----NKSVK-----KKESLSPINNKTKS 94		
Db	226	EITVSKHLMNIETNNKLPNTLIEIKYIYDEISKELNKLWLEDFPKKEXELS---NKISD 282		
Qy	95	YD-----IIPPSYSYR---NDKPNLSGENEDNSGN-----TNSNFPAN- 129		
Db	283	YDKKREQLSEYKMKLEIRNHYSQTNVNTKTEEEAKQYDKSNEHMTTPTIPTNEDEISKI 342		
Qy	130	TSBISIGCKD---NKQVTFIQKTHFLFACGIIKGRKS IKWICRENS-----EKITVCVCPDPRK 190		

Db 343 ISEVTKMDEILSKVNTYID-----FNKKYKETVNSEHSQFTLTDKIKAEVSDKE 393
Qy 181 IQLCVAMP-----LNSRLTMEKFEIFLSVNTAEKLYNKNVEGKDPST-FCNELRNS 233
Db 394 LKCEQSFNDKSLINETKNSIEKEYQ-----NINTLKKV-----DEVIKVCSTKES 441
Qy 234 FSDPRSSFIGDDMGNTDRVKGYINTKFSDDYKE--KNVEKLNLIKKEWKEK----- 285
Db 442 ITRFSSK-----QTILKMLNQNLIKVTKETNSIDKSYIEKFBQILT 482
Qy 286 -NKANLWNH-----IVNHKGNISKECAIIPAEPEQINLWIKEMN--ENFLMEKKRLFLNI 338
Db 483 GKQTKLENKFTFSLNNEHANNNE--LJKYFSDLKANLGINEENMLYNQFTEKEKTENDI 540
Qy 339 KDKCV-----ENKYEACFGGRLPCSSY-----TSFPMKSKTKTOMEVL 376
Db 541 KERNIHNEERISKIE-----IKIHASTYINISEETEREIGINISLNTKVFEKVK--ENV 592
Qy 377 TNLTKKNS-----GVDKNNFLNLPKNNK-----NDLDDPFKNEKE 414
Db 593 TNLNLIKELKHYDFSGKEGNTKYTDKIKKINDDIMAVSQIDQHINGLDIDQKSES 652
Qy 415 YDDLCDCRYTATIIFKSLFNGPAKND-----VDIASQINVNDLRGPGCNKSNNEKSNWCTG 470
Db 653 Y--VSEMKEQINKLEKYSNTEISNDNVEGIGKQKIIVTKI-----DKKNYIEE----- 700
Qy 471 TFTNKPCTCEPPRRQTLCLGRVTLHRGHEEDYKEHLLGASIEYEAQLLYKYEKEDENA 530
Db 701 --INKLLSEISKIEKNTSLKVKDINLSYQN-----LG-NLPLEQI--DEEKKAENT 750
Qy 531 LCSIIQNSYADLADIIGSDIIDKYGKMEENLAK--VNKKDKKNEPESLKIPE-----KW 585
Db 751 IKSM--EAYIDDLNIIK-----KKSQEIETEMDIKWDINKEMEALKISHDDDKC 798
Qy 586 WDEKENVVKVMSAVLKNKETCKD-YDKFOKIPQFLWFKEWGDDFCEKR-----KEKIY 639
Db 799 HDRSK-----NHKENISDIYDKSKIIQ-----DFSRESINDIKNKLQ 837
Qy 640 SPESFKVECKKDCDENTCKN-----KSEYKKWIDLKSEYEKQVDTKQKQK 690
Db 838 KNVS-----ESQVHNSDINQCLNEVANIYILKANKIKIID-KVKEYTSETEKKNQINDE 893
Qy 691 MYDNIDVKNKEANVYLKE-KSK-----EC-KDVFDDK-IFNESPN----- 729
Db 894 LNNSEKVIKIEGLSLKCRSKINSTLDDKDDIECKINVLKNTLNSETNITNFKN 953
Qy 730 -----EYDMCKKCDIEKYLNBIKYPKTKHDIYDITFSDTGDGTPISINANIN 779
Db 954 ABEYNKIVLSNFNNIEMADNKSQVILEIKKNGTND-HDY-----NIK 995
Qy 780 EQQSGKDTSTNGSETSDSPVSHPEPSDAINVEKLSGDESSHETRGILD-----IND 832
Db 996 ELKSHDKSGNYKTEAD-----QNKGAIQKNKELFEQYKEEVTLLNKYAVELKNK 1047
Qy 833 PSVTNN-----VNEVHDASNTQGSVNTSDITNGHSESSLNRTTNAQDIIKIGRSGEQSD 887
Db 1048 FDKTKNDKQIIEKIDA-----HNYCTLESKSEKKNWEIKN-EKIH-----ED 1092
Qy 888 NQENSHSSDNGSLTIGQVP-----SEDNTQNTYDSQPHRDTPNALASLPISDDKI 939
Db 1093 EVANNDSKNKAITSIKVSVEPFTKIKINEIRTKSD-DCLKETNDLEKQISNLSDTQE 1151
Qy 940 NEIEGPDSSRDSNGRG-----DTSNTNHDVRRNTNVSERR-----VNSHDFIRNGMANNA 991
Db 1152 TKL-----TENGKQLKTEELLESLEKQKKNIEQKKELDENVSK--IKN-IENTVN 1200
Qy 992 HHQVITQIENNGIIRGOEESAGNSVVKDNPKR-----SNFESSEN-----DHKKNI 1037
Db 1201 QHKQNYEI-----GIVEKINEIAKTNKQIESTKELIKETIOHIISSFNANDLEGIDSDENL 1257
Qy 1038 QEYNSRDKRVREIILKSKONKCN--BYSMEYCTYSDBERNSSPGPCSRERKKLCCQ 1094
Db 1258 GKYNT-BMGNIYEBFIK--SYNLITNVTVSKESITYNQIQNK-----RIDTQKEL--- 1306

Qy 1095 ISDYCLIKFNFYFSIYYNCIKS---BIKSPY-----KCFKSEGOSSIPYFAAGILVVIV 1147
Db 1307 -----LK-----NIENVNRAKSYLDYIKENEPRIVTHFPKKLNTVNDNF----- 1346
Qy 1148 LLLSSASRMCKSNEED-ICES--NIEATPEENNYLNKLSRIENOEOETNISDYSEYNY 1204
Db 1347 -----KNEYSKVNEDFNISNTVKNSTDENSLNLTQKEMYANIVNNTYYS-YKY 1400
Qy 1205 NEKNMY 1210
Db 1401 EAEINIF 1406
RESULT 37
G71609
hypothetical protein PF0650w - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 09-Jul-2004
C:Accession: G71609
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
Perteau, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.;
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600; MUID:99021743; PMID:9804551
A:Accession: G71609
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-2500 <GAR>
A:Cross-references: UNIPROT:O96223; UNIPARC:UPI0000078301; GB:AE001408; GB:AE001362; NID
A:Experimental source: clone 3D7
C:Genetics:
A:Gene: PF0650w

Query Match 5.1%; Score 328.5; DB 2; Length 2500;
Best Local Similarity 20.0%; Pred. No. 2e-06;
Matches 265; Conservative 178; Mismatches 463; Indels 421; Gaps 60;
Qy 18 NVIRINESIGR-----TLNQRQDESSDI---SRVNSPELNNHKTNIYD--SDY 62
Db 1375 NVVEKNSRFKKEHNTSMNLNVPNYENNRGKDI TNNNNISGDLVNGISTLSVKPYSTY 1434
Qy 63 EDVNN-----KLINSVENKSVKKESSLFINNKTYSYDIIPPSYVRNDKFNLSNED 117
Db 1435 NCISNIIBEKEIKKFNV---KKNSLNHINREKIY-----IGDNKKNYIENIC 1482
Qy 118 NSGNTSNNNF-ANTSISISGKONKQYTFIQKRPFLFACGIKRSIKWICRENSKIVCV 176
Db 1483 KCFPHSLGLNGGLPELVNKKKNLY---NNLHITNCLLFKK-----ETTEVLKKFL 1532
Qy 177 PDKIQLC--VANFLNSRLTMEKFEIF-----LISVNTAEKLY----- 215
Db 1533 PNNEINMSYVYNNHILYRLRMKNKYEDIITHDKLHVYHKLKELIKYQYKEYLLHKTVPYRN 1592
Qy 216 -----KNNEGKDSIFCNELNSFSDPRSSFIGDDMGNTDRVKGYINTKFSDYK 268
Db 1593 ICRNEHNNQKNDKTDIYINDNN-----KTELNIKISKENNEENKNTYMT---TSYK 1644
Qy 269 EKNVEKLNLIKKEWEKKNANLNHNMIVNHKGNISKECAIIPAEBPOI----- 316
Db 1645 ELLGNTVINFL-----DTFNLVDNIYSKEKYETDENLILANKKEPSISYFNFSNNYNDL 1697
Qy 317 -----NLWTIKWENENFLMEKKRLFLNIDKDC-----V 343
Db 1698 LKSDNVVEYIKDIYYDSYVDKNTYI-VYDKNVTFPKTNSFINDENGCVHLLTVPLEDEI 1756
Qy 344 ENKKEACGGCRLPCSSYTSFMKSKTKTQMEVLNTLYKKK-----NSGV--DKKNFLND-- 395
Db 1757 ENNNYEEKKKGHK-----RKIAHKNKMVNVLKRRKKIKYENENIISDKLVNMTYEY 1806
Qy 396 -----LFKQNNKNDLDDFFKNEKEYD-----DLCDCRYTATIIFKSL 432
Db 1807 NYTHKDEKEGSKCILKNKHNKLLKDKKLYNNVKKKWERMLPISKRRKLSQIR-- 1864

Qy 433 NGPARN-----DVIDASQINVDNRGFGCNYKSNNEK--SWNCTGTFTNKFP 477
Db 1865 NKIKKNQKSKILNKYKQVYSEFPFLSKKKNKNELLNGKHVEKKNKTALLNG-- 1922
Qy 478 GTCFPPRQTLCLGRYLLHGHEDYK-----EHLGASIVYEAQLLKYYKKEKDNALCS 533
Db 1923 -----GHTFVEDQKGEYKKEBERHIVQGEIKEKE--KYTLGGRGRSRRS 1967
Qy 534 IIONSY-----ADLADIIGSDIILKQYVKKMBENLNKVKDKRNEESLK----- 579
Db 1968 KESUSFRGRGSRSSKEVDTLKGR-----RDSLKGRKERSL-----KGRERDSLKGRKRD 2019
Qy 580 IFREKWDENKENYKVMASVAVLNKNETCK--DYDKFQKIPQFLRWFKEWG-----DDF 630
Db 2020 SFRGKERDSFRG-----KERETLKGREDSLKGREDSLKGREDSFRGKERDSF 2069
Qy 631 CEKKEKIIYSPESFKVECKKDC-----DENTCKKCE-YK-KWIDLKKSIEYKQVDKYT 684
Db 2070 RGERDSFRGKERDSFRGKERDTPFRSDRGSRFNKTGDVYKSRDINLYKEENKKDHY 2129
Qy 685 KDK-----NKKMYDNID-EVKNEANVYLKESKECKOV 717
Db 2130 VDKHYINKYPEKYSRKFNHSGSVHNAQKYDSLRYEKEPKYKITTENKKNGEVEI 2189
Qy 718 NFDDKIFNESPNYEDMCKCDEIKYLNKIKYPTKHDIDYDIDTFSDTFGDGTPIANAN 777
Db 2190 LKYSIENEKKNYD-----KEQNE-----NCILDKDT-----QCN 2220
Qy 778 INEQSGKDTNTGNSSETSDSPVSHESPESDAAINVEKLSDGESSETRGILINDPVTN 837
Db 2221 VNT-----KEKNLNDKKSFPSPNIK-----VKLEBEKSDDKR-----DDKND 2259
Qy 838 NVNVEDASNTQGSVNTSDITNGHSESSLNRTTNAQDIKIGRSGNSQSDNQENSHSSD 897
Db 2260 NTRKNLNDKKSFPSPN-----IKVLEBEKSDDKRDK----- 2294
Qy 898 NSGLITIGQVPESDNT--QNTYDSQNPDRTPNALASLPDDKI--NEIEGDFSRSRSEN 953
Db 2295 -----KNDNTRKNLNDK-----SFPSPNIKVKLEBEKSDDKRDK 2333
Qy 954 GRGDTTNTNHDVRR--TNI-VSRRNVNSHPDRNGMANNNAHQYITQIENN----- 1002
Db 2334 NDNTRKNLNDKKSFPSPNIKVKLEBEKSDDKRDKKNEN--TREKNLNDKCLP 2391
Qy 1003 -GIIRGOEESAGNSVYKPNKPNFNSENDHKKNIQVNSRDTKRVREELIKLS--KQN 1059
Db 2392 PSNIKVKLEBEKSDDKRDKDKDKDKDKDKDKDKDKDKDKDKDKDKDKDKDKKN 2451
Qy 1060 KCNN-BYSMEYCTYSDERNSSPGPCSRERKLCQISDYCLKYFNFYSIRVYNCIKSEI 1118
Db 2452 KQNNKVKKEKNIKNDKN-----NKNKKVA-----KQNEIKKEIKDEI 2492
Qy 1119 KSPRYK 1125
Db 2493 KDNKKC 2499

RESULT 38
A45605
A;Title: parasite-infected erythrocyte surface antigen MESA - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A45605; A54517
R;Coppel, R.L.
Mol. Biochem. Parasitol. 50, 335-347, 1992
A;Title: Repeat structures in a Plasmodium falciparum protein (MESA) that binds human erythrocyte surface antigen
A;Reference number: A45605; MUID:92158014; PMID:1741020
A;Accession: A45605
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1526 <COP>
A;Cross-references: UNIPROT:P08116; UNIPARC:UPI0000177P9B

A;Experimental source: Papua New Guinean isolate FC27
A;Note: sequence extracted from NCBI backbone (NCBIN:83648, NCBI:83656)
R;Coppel, R.L.; Culvenor, J.G.; Bianco, A.E.; Crewther, P.E.; Stahl, H.D.; Brown, G.V.; Mol. Biochem. Parasitol. 20, 265-277, 1986
A;Title: Variable antigen associated with the surface of erythrocytes infected with mature Plasmodium falciparum
A;Reference number: A54517; MUID:87014571; PMID:3531849
A;Accession: A54517
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 222-443, 'K' 445 <COP>
A;Cross-references: UNIPARC:UPI000002C732; GB:M15319; NID:g160060; PID:g552170
C;Superfamily: ring-infected erythrocyte surface antigen; dna amino-terminal homology
C;Keywords: surface antigen; tandem repeat

Query Match 5.1%; Score 328; DB 2; Length 1526;
Best Local Similarity 20.08; Pred. No. 1.2e-06;
Matches 270; Conservative 203; Mismatches 474; Indels 402; Gaps 59;

Qy 46 SPELNHNHTNIYDSYEDVNNKLINSFVENKSVKKRSLSFINNKTYSVDIIPPSYSYR 105
Db 316 SKETGESKETRIYE--ETKYNKITSEFRETEVNVK-----ITEESKDRGNKVSGPYE 365
Qy 106 N-----DKFNSLSEHEDNSGNTNSNFFANTSEISGKDNKQVTFIQKTHLPACG 155
Db 366 NSENSNVTSSEETKLAKEENEGEKLGEN-VNDGASENSEDPKLQTEQEN-----G 418
Qy 156 IKRSIKW--ICRENSEKITVCPDRK-----IQLCVANFL-NSRLETME----- 197
Db 419 TKESSEETKDDKPRENEKADNKKSKKKKSPFQMLGCMFLCNKNIETDDEETLVVKD 478
Qy 198 --KFKBIFLSVNTAEKLLYNQNEGKDPISFCNELRNSFSDFRSSFIDGDDMDFGNTDRV 255
Db 479 DAKKKHFLREANTEK-----NDNEKKDKLL--GS-----GDKEDVKEKDEQ 519
Qy 256 KG-YINTKFDYDYKKNVEKLNLI-----KKEWEKKKANLWNMIVNHKGNISKECAII 309
Db 520 KDKVLSEGDVEDKERNDEQDKVLGDEKDEKERN-----GKKDKVI 564
Qy 310 PAEPQINLWIKENNFLEKELFLNIDKQCVENKYEACFGGRLPCSSYTSFMMKS 369
Db 565 GSEKIQ-----KEIKEK--VEKR-----VKKCKCKVK-----KGIKEN 596
Qy 370 KTO-----MEVLNLYKKNQSGVDKNN-----FLNDLPKKNQNDLD 406
Db 597 DTEGNDKVGPEIIEVEBEIKKQVEDGKENDTEGNDKVGPEIIEVEBEIKKQVE 656
Qy 407 DFPK-NEKEVDD-----LCDCRYTATIIISFLNGPAKNDVD-----IASQINVN 449
Db 657 DGIKENDTEGNDKVGPEIIEVEBEIKKQVEBEIKKQVEBEIKKQVEBEIIEVEKBE 716
Qy 450 DLRGFGCNYKSNNEKSNW-----CTGTFTNKFPGTCEPRROT-----LCIGRTYLL 496
Db 717 IKQVEGKENDTESKDKVIGQBIIEVEKBEIKKQVEBEIKKQVEBEIKKQVEBEIIEVE 776
Qy 497 -----HRGHEEDY--KEHLGASIVYEAQLLKYYKKEKDNALCSIIQNSYADL 542
Db 777 EEVKVKVKKVEKIKENHTESKDKVIGQBIIEVEV-----KBEIKQVEBEIKENDTES 831
Qy 543 ADIIGSDIILKQYVKKMBENLNKVKND-----KKNNEESLKIIFREKWDENKEN 592
Db 832 KDKVIGSEVIGKGVNBEENEGENKDKVTQKEKVKVEKVKKKVKRKRNNK--NERKDN 889
Qy 593 V--WKVMSAVLNKNETCKEDYKQKIPQFLRWFKEWGDDDFCEKKEKIYSPESFKVECKK 650
Db 890 VIGKEIMKEDVNEKDTA--NKKKBIEQ-----EKEKEVKEKEVEKEVEE 933
Qy 651 KDCDENTCKKCKEYKWIIDLKSEYEKQVDKYTKDK--NKKMYDNIDVKNKEA----- 703
Db 934 KEKEE--VKEK-EVKEKEVEKEVEKEVEKEKTESKDEIKBEKEVEKEVEKEDTENKD 990
Qy 704 ----NVYLKESKEC-----KDNVFD-----KIFNESPNYE----- 732
Db 991 KVIGQBIIEBEIKVKRKRNNKKNENKDNVIGQIMNEDVNEKDTANKDKVIEQKE 1050

Search completed: November 21, 2005, 20:34:43
Job time : 90 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2005, 20:34:49 ; Search time 179 seconds
(without alignments)
2824.433 Million cell updates/sec

Title: US-10-677-980-2
Perfect score: 6481
Sequence: 1 MKGYFNIYFLIPLIFLYNYI.....VQETNISDYSEYNNKKNMY 1210

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pcp.*
3: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pcp.*
4: /cgn2_6/ptodata/1/pubppaa/US10A_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pcp.*
6: /cgn2_6/ptodata/1/pubppaa/US11_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
2	6481	100.0	1210	5	US-10-677-980-2
3	6081	93.8	1143	3	US-09-924-154-14
4	1557	24.0	1435	4	US-10-153-273-4
5	1537.5	23.7	1421	3	US-09-924-154-13
6	1200.5	18.5	616	4	US-10-293-913A-4
7	1199.5	18.5	616	4	US-10-293-913A-2
8	1112	17.2	1086	3	US-09-924-154-15
9	1088	16.8	1568	5	US-10-712-533A-12
10	993.5	15.3	1501	3	US-09-924-154-17
11	761.5	11.7	749	4	US-10-153-273-6
12	754	11.6	1115	4	US-10-153-273-2
13	750.5	11.6	972	3	US-09-924-154-16
14	601.5	9.3	3542	4	US-10-087-013-2
15	495.5	7.6	2710	4	US-10-153-273-12
16	411	6.3	1985	5	US-10-732-923-3351
17	382	5.9	6761	5	US-10-732-923-15035
18	370.5	5.7	1527	5	US-10-732-923-3354
19	363	5.6	407	4	US-10-087-013-8
20	360.5	5.6	921	4	US-10-153-273-8
21	353	5.4	4226	5	US-10-732-923-22586
22	352	5.4	4226	5	US-10-732-923-22707
23	349.5	5.4	2548	5	US-10-732-923-15009
24	348	5.4	294	4	US-10-087-013-7
25	338	5.2	2719	5	US-10-732-923-8668
26	334.5	5.2	2133	5	US-10-732-923-15030
27	333.5	5.1	351	4	US-10-087-013-9

28	331.5	5.1	2910	5	US-10-732-923-3342	Sequence 3342, Ap
29	329	5.1	3124	5	US-10-732-923-22709	Sequence 22709, A
30	327.5	5.1	308	4	US-10-087-013-11	Sequence 11, Appl
31	326.5	5.0	3127	5	US-10-732-923-22588	Sequence 22588, A
32	325	5.0	700	4	US-10-153-273-10	Sequence 10, Appl
33	323.5	5.0	1478	5	US-10-732-923-3353	Sequence 3353, Ap
34	319.5	4.9	311	4	US-10-087-013-10	Sequence 10, Appl
35	318.5	4.9	1939	5	US-10-732-923-3340	Sequence 3340, Ap
36	318	4.9	2110	5	US-10-732-923-8666	Sequence 8666, Ap
37	310.5	4.8	1441	5	US-10-732-923-3352	Sequence 3352, Ap
38	308	4.8	2184	4	US-10-304-095-6	Sequence 6, Appl
39	301.5	4.7	277	4	US-10-153-273-15	Sequence 15, Appl
40	295	4.6	1647	5	US-10-732-923-8314	Sequence 8314, Ap
41	293.5	4.5	4688	4	US-10-282-122A-76865	Sequence 76865, A
42	289.5	4.5	861	3	US-09-820-843A-109	Sequence 109, App
43	287	4.4	1881	4	US-10-032-585-7646	Sequence 7646, Ap
44	285.5	4.4	282	4	US-10-153-273-16	Sequence 16, Appl
45	285.5	4.4	980	5	US-10-732-923-3343	Sequence 3343, Ap

ALIGNMENTS

RESULT 1
US-10-677-980-2
; Sequence 2, Application US/10677980
; Publication No. US20050239730A1
; GENERAL INFORMATION:
; APPLICANT: Mayer, D.C.Ghislaire
; APPLICANT: Mailer, Louis H.
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ERYTHROCYTE
; FILE REFERENCE: NIH209 001C1
; CURRENT APPLICATION NUMBER: US/10/677,980
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: PCT/US02/10071
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: US 60/281130
; PRIOR FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1210
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-677-980-2

Query Match	100.0%;	Score 6481;	DB 5;	Length 1210;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1210;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	MKGYNFYFLIPLIFLYNVIRINESIIGRTLYNRQDESSDISRVNSPELNNHNTIYDS	60	
Db	1	MKGYNFYFLIPLIFLYNVIRINESIIGRTLYNRQDESSDISRVNSPELNNHNTIYDS	60	
Qy	61	DYEDVNNKLNFSVFNKSVKKSLSFINNKTYSYDIIPPSYSYRNDKFNLSNENSG	120	
Db	61	DYEDVNNKLNFSVFNKSVKKSLSFINNKTYSYDIIPPSYSYRNDKFNLSNENSG	120	
Qy	121	NTNSNFPANTSEISIGKDNQYTFIQKTHLFAAGIKKRSIKWICRENSKITVCPDRK	180	
Db	121	NTNSNFPANTSEISIGKDNQYTFIQKTHLFAAGIKKRSIKWICRENSKITVCPDRK	180	
Qy	181	IQLCVANFLNRLSEIETMEKFEIPLISVNTAKLLYNKNEKQPSIFCNELRNSFDRSS	240	
Db	181	IQLCVANFLNRLSEIETMEKFEIPLISVNTAKLLYNKNEKQPSIFCNELRNSFDRSS	240	
Qy	241	FIGDDMDFGNTDRVKGYINTKESDYKKEVKEKLANIKKEWEKKNKANLWNHMTVNHKG	300	
Db	241	FIGDDMDFGNTDRVKGYINTKESDYKKEVKEKLANIKKEWEKKNKANLWNHMTVNHKG	300	
Qy	301	NISKECAIIPAEBPQNLWIKWENFLMEKKLFLNIDKQCVENKKEACFGGCRLPDS	360	
Db	301	NISKECAIIPAEBPQNLWIKWENFLMEKKLFLNIDKQCVENKKEACFGGCRLPDS	360	

Db 301 NISKECAIIPAEPOINLWIKENENFLMEKKRLFLNIDKCVENKKYKACFGCRLPCS 360
Qy 361 SYTSFMKSKTQMEVLNLYKKNSGVDKNNFLNDFKNNKNDLDDFFKNEKEYDDLCD 420
Db 361 SYTSFMKSKTQMEVLNLYKKNSGVDKNNFLNDFKNNKNDLDDFFKNEKEYDDLCD 420
Qy 421 CRYTATIILKSLFNGPAKNDVDIASQINVDLGRFGCNYSNNKSWNCTGTFNKPFGTC 480
Db 421 CRYTATIILKSLFNGPAKNDVDIASQINVDLGRFGCNYSNNKSWNCTGTFNKPFGTC 480
Qy 481 EPPRQTLCLGRVYLLHRGHEEDYKEHLLGASIEAQLLYKKEKDNALCSIIQNSYA 540
Db 481 EPPRQTLCLGRVYLLHRGHEEDYKEHLLGASIEAQLLYKKEKDNALCSIIQNSYA 540
Qy 541 DLADIILKSGDIIKDYGGKMEENLKNVNDKRRNEESLKIIFREKWDENKENVKWSAV 600
Db 541 DLADIILKSGDIIKDYGGKMEENLKNVNDKRRNEESLKIIFREKWDENKENVKWSAV 600
Qy 601 LKNETCKDYDKFKIPOFLRWPKEWGDGDFCEKKEKIKYSPESPKVECKKDCDENTCKN 660
Db 601 LKNETCKDYDKFKIPOFLRWPKEWGDGDFCEKKEKIKYSPESPKVECKKDCDENTCKN 660
Qy 661 KCSEYKKWIDLKSEYKQVDKYTKDNKKMYDNIDEVKNKEANVYLKESKECKOVNFD 720
Db 661 KCSEYKKWIDLKSEYKQVDKYTKDNKKMYDNIDEVKNKEANVYLKESKECKOVNFD 720
Qy 721 DKIFNESPNEYEDMCKKDEIKYLNEIKYPTKHDIDYDIDTFSDFGDTGPISINANINE 780
Db 721 DKIFNESPNEYEDMCKKDEIKYLNEIKYPTKHDIDYDIDTFSDFGDTGPISINANINE 780
Qy 781 QOSGKOTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSESTRGILDINDPSVTNNV 840
Db 781 QOSGKOTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSESTRGILDINDPSVTNNV 840
Qy 841 EVHDASNTQGSVNTSDITNGHSESSLRNTTNAODIKIGRSGNQSDNOENSHSSDNG 900
Db 841 EVHDASNTQGSVNTSDITNGHSESSLRNTTNAODIKIGRSGNQSDNOENSHSSDNG 900
Qy 901 SLTIGQVPSDNTQNTYDSQNPHRDTPNALASLPSDDKINEIEGDFSRSSENGRDTTS 960
Db 901 SLTIGQVPSDNTQNTYDSQNPHRDTPNALASLPSDDKINEIEGDFSRSSENGRDTTS 960
Qy 961 NTHDVRRTNIVSERVNSHDFIRGMANNNAHQYITQIENNGIIRGOESAGNSVNYKD 1020
Db 961 NTHDVRRTNIVSERVNSHDFIRGMANNNAHQYITQIENNGIIRGOESAGNSVNYKD 1020
Qy 1021 NPKESTFSSNDHKNIQEVNSRDTKRVREBIIKLSKONKCNBYSMEYCTYSDERNSSP 1080
Db 1021 NPKESTFSSNDHKNIQEVNSRDTKRVREBIIKLSKONKCNBYSMEYCTYSDERNSSP 1080
Qy 1081 GPCSRERBKLCQOISDYCLKYNFYNSIYNYCNICKSEIKSPYKCFKSEGOSSIPYFAAG 1140
Db 1081 GPCSRERBKLCQOISDYCLKYNFYNSIYNYCNICKSEIKSPYKCFKSEGOSSIPYFAAG 1140
Qy 1141 GILVIVILLSSARMGKSNEYDIGESNIEATFEENNYLKLRIFNQEVQETNIDYS 1200
Db 1141 GILVIVILLSSARMGKSNEYDIGESNIEATFEENNYLKLRIFNQEVQETNIDYS 1200
Qy 1201 EYNYNEKMY 1210
Db 1201 EYNYNEKMY 1210

RESULT 2

US-09-924-154-14
; Sequence 14. Application US/09924154
; Patent No. US20020127241A1
; GENERAL INFORMATION:
; APPLICANT: Narum, David L.
; APPLICANT: Sim, Kim L.
; TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use
; FILE REFERENCE: 05213-0465 43170-262105
; CURRENT APPLICATION NUMBER: US/09/924,154

; CURRENT FILING DATE: 2001-08-07
; PRIOR APPLICATION NUMBER: US 60/223,525
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 1143
; TYPE: PRT
; ORGANISM: Mammalian
US-09-924-154-14

Query Match 93.8%; Score 6081; DB 3; Length 1143;

Best Local Similarity 99.8%; Pred. No. 0;
Matches 1130; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MKGYFNIYFLIPLIFLYNVIRINESIIGRTLYNRQDESSDISRVNSPELNNHKTNIYDS 60
Db 7 MKGYFNIYFLIPLIFLYNVIRINESIIGRTLYNRQDESSDISRVNSPELNNHKTNIYDS 66
Qy 61 DYEDVNNKLINSFVENKSVKKGSLSFNNKTSYDIIIPPSYSYRNDKFNLSSENDSNG 120
Db 67 DYEDVNNKLINSFVENKSVKKGSLSFNNKTSYDIIIPPSYSYRNDKFNLSSENDSNG 126
Qy 121 NTNSNFPANTSEISIGKDNKQYTFIOKRTHLFACGIRKRSIKWICRENSKITTVCVPDRK 180
Db 127 NTNSNFPANTSEISIGKDNKQYTFIOKRTHLFACGIRKRSIKWICRENSKITTVCVPDRK 186
Qy 181 IQLCVAFNLSRLSETEWKEFKEIFLISVNTAEKLLYNKEGKDPISIFCNELRNSPSPDRSS 240
Db 187 IQLCVAFNLSRLSETEWKEFKEIFLISVNTAEKLLYNKEGKDPISIFCNELRNSPSPDRSS 246
Qy 241 FIGDDMDFGNTDRVKGYNINTEKPSDYKKEKNNVEKLANNIKKEWKEKNNKANNLWNNHMI 300
Db 247 FIGDDMDFGNTDRVKGYNINTEKPSDYKKEKNNVEKLANNIKKEWKEKNNKANNLWNNHMI 306
Qy 301 NISKECAIIPAEPOINLWIKENENFLMEKKRLFLNIDKCVENKKYKACFGCRLPCS 360
Db 307 NISKECAIIPAEPOINLWIKENENFLMEKKRLFLNIDKCVENKKYKACFGCRLPCS 366
Qy 361 SYTSFMKSKTQMEVLNLYKKNSGVDKNNFLNDFKNNKNDLDDFFKNEKEYDDLCD 420
Db 367 SYTSFMKSKTQMEVLNLYKKNSGVDKNNFLNDFKNNKNDLDDFFKNEKEYDDLCD 426
Qy 421 CRYTATIILKSLFNGPAKNDVDIASQINVDLGRFGCNYSNNKSWNCTGTFNKPFGTC 480
Db 427 CRYTATIILKSLFNGPAKNDVDIASQINVDLGRFGCNYSNNKSWNCTGTFNKPFGTC 486
Qy 481 EPPRQTLCLGRVYLLHRGHEEDYKEHLLGASIEAQLLYKKEKDNALCSIIQNSYA 540
Db 487 EPPRQTLCLGRVYLLHRGHEEDYKEHLLGASIEAQLLYKKEKDNALCSIIQNSYA 546
Qy 541 DLADIILKSGDIIKDYGGKMEENLKNVNDKRRNEESLKIIFREKWDENKENVKWSAV 600
Db 547 DLADIILKSGDIIKDYGGKMEENLKNVNDKRRNEESLKIIFREKWDENKENVKWSAV 606
Qy 601 LKNETCKDYDKFKIPOFLRWPKEWGDGDFCEKKEKIKYSPESPKVECKKDCDENTCKN 660
Db 607 LKNETCKDYDKFKIPOFLRWPKEWGDGDFCEKKEKIKYSPESPKVECKKDCDENTCKN 666
Qy 661 KCSEYKKWIDLKSEYKQVDKYTKDNKKMYDNIDEVKNKEANVYLKESKECKOVNFD 720
Db 667 KCSEYKKWIDLKSEYKQVDKYTKDNKKMYDNIDEVKNKEANVYLKESKECKOVNFD 726
Qy 721 DKIFNESPNEYEDMCKKDEIKYLNEIKYPTKHDIDYDIDTFSDFGDTGPISINANINE 780
Db 727 DKIFNESPNEYEDMCKKDEIKYLNEIKYPTKHDIDYDIDTFSDFGDTGPISINANINE 786
Qy 781 QOSGKOTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSESTRGILDINDPSVTNNV 840
Db 787 QOSGKOTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSESTRGILDINDPSVTNNV 846
Qy 841 EVHDASNTQGSVNTSDITNGHSESSLRNTTNAODIKIGRSGNQSDNOENSHSSDNG 900


```

Db      985  EEVKEHTSNSDNVQSGGIYVMNVYEKELKOTLEPNFSSSLDEGKAHELSSEPNLSSDQMS 1044
Qy      847  NTQGSVNTSDITN---GHSESSLNR-----TTNAQDIKIGRSGNEQSDN---QENS 892
Db      1045  NTPGLDNTSBETTERISNNEYKVNEREDERTLTKEVEDIVLKGSHMNRSDGDELYDENS 1104
Qy      893  SHSDNSGSLTI-CQVSEDTQNTY-----DSQPHRDTNALA-----SLP- 934
Db      1105  DLSTVNDESEDAEAKMGNDTSEMHSNOSHIESDQDQNDKMTVGDGLTTHVQNEISVPV 1164
Qy      935  ---SDDKINE-----IEGFDSSRDSNGRGDTTSTNTHDVR-----R 967
Db      1165  TGEIDEKLRESKESIKHABEERLSHTDIHKINPDSNLTSLHLKDIREENERHLTNQN 1224
Qy      968  TNIVSRVNSHDPIRNGMANNNAHQYITQ-----IENNGIIRQGESAGSNVYKDN- 1021
Db      1225  INISOERDLQKHGF---HTMNNLHGDGVBSERSQINSHSHGNRQDRGSGNVLNMRSSNN 1280
Qy      1022  -----PKRSNFSSNDHKNIQBYNSRDTKRVREBIIKLSKQKCNMEYMEYCTYSDE 1075
Db      1281  NNFNIPESRYNL-----YDKKLDLDIYENRNDSTTKELIKKLAETIKCENBISVKYCDHMIH 1337
Qy      1076  RNSSPGSCSRERKLCACOSDYCLKYFNFYISIEYVNCIKSEIKSEIKSEIKSEIKSEIKSE 1134
Db      1338  BEPLKCTKEKTRNLCAVSDYCMSTFTYDSEBYNCTKREPDPSYTCRKEAPSSM 1396

RESULT 4
US-09-924-154-13
; Sequence 13, Application US/09924154
; Patent No. US20020127241A1
; GENERAL INFORMATION:
; APPLICANT: Narum, David L.
; APPLICANT: Sim, Kim L.
; TITLE OF INVENTION: Anti-plasmodium Compositions and Methods of Use
; FILE REFERENCE: 05213-0465 43170-262105
; CURRENT APPLICATION NUMBER: US/09/924,154
; CURRENT FILING DATE: 2001-08-07
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 1421
; TYPE: PRT
; ORGANISM: Mammalian
US-09-924-154-13

```

```

Query Match      23.7%; Score 1537.5; DB 3; Length 1421;
Best Local Similarity 28.1%; Pred. No. 2.4e-77;
Matches 411; Conservative 211; Mismatches 467; Indels 375; Gaps 44;

Qy      1  MKGFNIYFLIPLIFLYNVIRINESIIGRTLYNRQDESSDISRVNSPELNNHKNITVDS 60
Db      1  MKCNISYFFASFFVLVFAKARNEYDI-----KENEKFLDVYKEKFNELDKKKYGVNQKT 55
Qy      61  DYEDVNNKLINSFVENKSVKKRSLSPINN-----KTKSY---DIIPPSYSRNDKFNSL 112
Db      56  D-----KKIIFTIENK-----LDILNNSKFNKWSYGTDPNDIKNMSLIN----- 96
Qy      113  SENEDNSGNTNNSNFPANTSISIGKDNKQ-----YTFIQKTH----- 150
Db      97  KHNEEMFNNNYQSFLLSTS---SLIKONKYVPINAVRVSRLSFLDSRINNNGRNTSSNVEY 154
Qy      151  LFACIGIKRSIKWICRENSEKIT-VCPVDRKIQLCVANFLNLSRLTWEKFEIPLISVNT 209
Db      155  LSNCREKRRKMKWCKKNDKNSRNTVCIPDRRIQICVNLISIIKIYTKETWKDHFIEASKK 214
Qy      210  EAKLLYNKNEGKOPSIFCNELRNSPDRSSSFGIDMDFGGNTDRVRYGINTKFSDYKXE 269
Db      215  ESQLLKKNNDKNYSKFCNDLKNFLDYGHLAWGNDMDFGGYSTKAENKIQEVFKGAHGE 274
Qy      270  KNVEKLNNIKKEWKEKKNANLWNEHIVNHKGNISKECALIPAEEFPQINLMIKENWENFLM 329

```

```

Db      275  ISBHKIKNFRKKNWNEFREKLWEAMLSEHKNNIN-NCKNIPQBELQITOWIKEMHGSEFL 333
Qy      330  EKRLPLNTKDKVENKYEACFGGRLPCSSYTSFMKSKTQMEVLNTLYKKKNQSVDK 389
Db      334  ERNRSKLPKSKCKNNTLYEACEKECIDPCMKYRDMIIIRSKFPHWTLSEKEYEQK--VPK 391
Qy      390  NNFLNDLPK-KNNKND--LDDFFKN-EKEYDDLCDCRYTATIIKSFNGPAKNDVDIASQ 445
Db      392  ENAENYLIKISEKNDAKVSLLNNDARYSKYCDCKHKTTLTKVSLVNGNDNTIKEKREH 451
Qy      446  INVNDLRGGCNYKS--NNEKSNWCTGTFTNKEPGTCEPRRQTLCLGRYTLHRGHEED 503
Db      452  IDLDDFSKGCNKSVDVTNTTKWCEKKPYKSLTKDVCVPPRQBELCLGNTDRIYKLLM 511
Qy      504  YKEHLGASIEAQLKYKKEKENALCSIIQNSVADLADIIGSDIIKDYVKGKMEEN 563
Db      512  IKEHILAIYESRIILKRYKKNDDKEVCKLINKTADIRDIIGTDYWNDSNRKLVGK 571
Qy      564  LNKVNKDKKRNEESLKIIFREKMWDEKENYKVSALVKNKETCKDYDKPKIQPQFURWF 623
Db      572  INTNSNVHRNKQNDKLFREDWVKVICKDVNVVISWVFXDKTVCKE--DIENIPQPFRRWF 630
Qy      624  KEGWDDFCERKKEKIYSPSPFKVECKKDCDENTCKNKSEYKWKIDLKSEYEKQV--- 680
Db      631  SEWGDDYCDQDKTMI---ETLKVECKEKPCEDDNCKRCKNSYKEWISKKEEYNKQAKQY 687
Qy      681  DKYTKDNKKNKYDNIDEVKNKEANVYLKESKCKDVNFDDKIFNESPNEYEDMCKKDE 740
Db      688  OEYQKGNKYKY---SEFKSIKPEVYLKYSEKSNLNFEDFEKELHSDYKKNCTWCPE 744
Qy      741  IKYL-----NEIKYP----- 750
Db      745  VKDVPISIIIRNNRQTSQEAVPESRTEIAHRTETRTDERKNOEPANKDLKPQSVSGENG 804
Qy      751  -----KTKHD 755
Db      805  KDLQBEDLGGSRSEDEVTQFVGNHGIKGBEDQTLGKSDAIPNIGEPGTGISTTEESRHE 864
Qy      756  -----IYDIDTFSDTF-----GDGTPISINAMINEQOSG-KDTSNT----- 790
Db      865  EGHKQALSTSVDEPELSDTLQHEBTKENDKLPLESSTITSSTESGSSDTTEPSPISSE 924
Qy      791  --GN-----SETSDSPVSHPEPSDA--AINVEKLSGD----- 818
Db      925  PKGNEQKRRDDDSLKSISVSPENSRPETDAKTNLLKLKGDVDISMPKAVIGSSPNDNI 984
Qy      819  -----ESSESTRGILDIN----- 831
Db      985  NVTEQDNTSGVNSKPLSDDDVRPDKNHEEYKHEHTSNSDNVQSGGIYVMNVYEKELKDTLE 1044
Qy      832  DPSVTNNVNEVH-----DASNTQGSVNTSDITN---GHSESSLNR----- 869
Db      1045  NPSSLDGKAHELSSEPNLSSDQMSNTPGLDNTSEETTERISNNEYKKNREGERRTL 1104
Qy      870  TTNAQDIKIGRSGNEQSDN---QENSNSHSDNSGSLTI-CQVSEDTQNTY----- 917
Db      1105  TKEYEDIVLKSHMNRSDGELYDENSIDLSTVNDESEDAEAKMGNDTSEMHSNOSHIE 1164
Qy      918  -DSQPHRDTNALA-----SLP-----SDDKINE-----IEGFDSSRDS 951
Db      1165  SDQQNDKMTVGLGTTHVQNEISVPVTEIGDEKLRESKESIKHABEERLSHTDIHKIN 1224
Qy      952  ENRGDITSTNTHDVR-----RTNIVSRVNSHDPIRNGMANNNAHQYITQ-- 998
Db      1225  PEDRNSNTLHLKDIREENERHLTNQNINISOERDLQKHGF---HTMNNLHGDGVBSERS 1280
Qy      999  ---IENNGIIRQGESAGSNVYKDN-----PKRSNFSSNDHKNIQBYNSRDTKRV 1048
Db      1281  QINHSHHGNRQDRGSGNVLNMRNNSNNNFNIPSRYNL---YDKKLDLDIYENRNDSTT 1337
Qy      1049  REBIIKLSKONKCNMEYMEYCTYSDBERNSSPGCSREERKLCACOSDYCLKYFNFYSI 1108

```

Db 1338 KELIKLARIKNCENISVYCDHMIHBEIPLKCTCKEKNLCAVSDYCMVSYTYDSE 1397
Qy 1109 EYINCISEIKSPYKCFKSEGS 1132
Db 1398 EYINCTKEFDPSYTCFRKEAPS 1421

RESULT 5
US-10-293-913A-4
; Sequence 4, Application US/10293913A
; Publication No. US20040022805A1
; GENERAL INFORMATION:
; APPLICANT: Narum, David
; APPLICANT: Liang, Hong
; APPLICANT: Fuhrmann, Steve
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Synthetic Genes for Malarial Proteins and Methods of Use
; FILE REFERENCE: 05213-0464 (43170-280206)
; CURRENT APPLICATION NUMBER: US/10/293,913A
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/345,051
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 4
; LENGTH: 616
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic EBA-175 RII
US-10-293-913A-4

Query Match 18.5%; Score 1200.5; DB 4; Length 616;
Best Local Similarity 38.4%; Pred. No. 6.3e-59;
Matches 246; Conservative 111; Mismatches 236; Indels 47; Gaps 13;

Qy 154 CGIKRSIKWT CRENSEKIT--VCVPDRKIQLCVANFLNSRLTMEKFKEIFLISVNTAK 212
Db 14 CREKRGKMKWCKKKNDNSNVCIIPDRRIQLCIVNLAIKTYTETWKDHFIEASKESQ 73

Qy 213 LLYNKGEGDSIECNELNSPSPERSPIGDDMDFGNTDRVKGYNITKFSDYKKNV 272
Db 74 LLLKKNKNYSKFCNDLNSFLDYGLAMGNDMDFGYSTKAENKIQEVFKGAGHISE 133

Qy 273 EKLNIKKWKKKKNKANIWNHMI VNHKGNKECAIIPAEPPQINLWIKENWENFLMEKK 332
Db 134 HKIKNFKKWNWREKLEWMLSEHKNNIN--NKNIPQEBLQITQWIKWHGFLERD 192

Qy 333 RLFLNIKDKCVENKKEACFCGCRPLCSSYTSFMKKSQTQMEVLTNLYKKNSGVDKNNF 392
Db 193 NRAKLPSKCKNNALYEAKECIDPCMKYRDWIIRSKFEWHTLSKEYETQK--VPKENA 250

Qy 393 LNDLFK--KNKND--LDDFFKN--EKEYDDLCRYTATIIKSPNGPAKNDVDIASQINV 448
Db 251 ENYLKISEKNDKAVSLLNLCDAEYSKYCDCKHTTLVKSVLGNDNTIKREHIDL 310

Qy 449 NDLRGFGNYKS--NNEKSNWCTGTFTNKPFGTCPPRQTLCLGRVTLHRGHEEDYKE 506
Db 311 DDFKFGCKNSVDNTNKTWCKPKYKSTKDVCPVPRQELCIGNDRIYDKNLLMIKE 370

Qy 507 HLLGASIEAQLLYKYKEDENALCSIIQNSYADLADIIKGSDDIIKDYKQKMEENLNK 566
Db 371 HILAIAIESRLKRYKNDKCKVCKIIQKTFADIRDIIGTGYNDLSNRKLVGKINT 430

Qy 567 VNKDKRNEESLKI PREKWDENKENVKMSAVLNKNCETKQYDKFKQIPQIRWPKW 626
Db 431 NSNTVHRNKQNDKLFREDWVWVIKKDVVNVISWVFKDVTCKE--DDIENIPQFRWSEW 489

Qy 627 GDDCEKEKEKIIYSFESKCKKCDENTCKNCKSEYKWKIDLKSEYKQV---DKY 683
Db 490 GDDYCDQKTKMI---ETLKVECKBPCEDDKCKRCKNSYKEMISKCKEYNNKQAKQY 546

Qy 684 TKDKNKKMYNDIDBVKNEKNAVYLKESKCKDVNFDDKIFNESPNEVEDNCKCKDEIKY 743

Db 547 QKGNRYKMY---SEFKSIKPEVLYLKYSKCSNLNFEDEPKELHSDYKNCCTMCPEVK- 602
Qy 744 LNEIKYPTKTHDIDYDITDFTSDGTGTPISINANINQOS 783
Db 603 -----DVPISIRN-NEQTS 616

RESULT 6
US-10-293-913A-2
; Sequence 2, Application US/10293913A
; Publication No. US20040022805A1
; GENERAL INFORMATION:
; APPLICANT: Narum, David
; APPLICANT: Liang, Hong
; APPLICANT: Fuhrmann, Steve
; APPLICANT: Sim, B. Kim Lee
; TITLE OF INVENTION: Synthetic Genes for Malarial Proteins and Methods of Use
; FILE REFERENCE: 05213-0464 (43170-280206)
; CURRENT APPLICATION NUMBER: US/10/293,913A
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/345,051
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2
; LENGTH: 616
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic EBA-175 RII
US-10-293-913A-2

Query Match 18.5%; Score 1199.5; DB 4; Length 616;
Best Local Similarity 37.4%; Pred. No. 7.2e-59;
Matches 252; Conservative 113; Mismatches 239; Indels 69; Gaps 15;

Qy 121 NTNSNNTANTSEISIGKNDKQYTFIQKTHLPACGIRKRSIKWICRENSKIT--VCVPDR 179
Db 3 NTSSNN-----EV-----LSNCREKRGKMKWCKKKNDNSNVCIIPDR 40

Qy 180 KIQLCVANFLNSRLTMEKFKEIFLISVNTAKLLYNKNEGKOPSIFCNELNSPSPERS 239
Db 41 RIQLCIVNLISIIKTYTETWKDHFIEASKESQLLKKNKNYSKFCNDLNSFLDYGH 100

Qy 240 SPIGDDMDFGNTDRVKGYNITKFSDYKKNVLEKNNIKKEWKKKANIWNHMI VNHK 299
Db 101 LAMGNDMDFGYSTKAENKIQEVFKGAGHISEHKIKNFKCKWNEFREKLWEAMLSEHK 160

Qy 300 GNISKECAIIPAEPPQINLWIKENWENFLMEKKRLFLNIKDKCVENKKEACFCGCRPLC 359
Db 161 NNIN--NKNIPQEBLQITQWIKWHGFLERDNRKLPKSKCKNNLTLYEAKECIDPC 219

Qy 360 SSYTSFMKKSQTQMEVLTNLYKKNSGVDKNNFLNLFK--KNKND--LDDFFKN--EKEY 415
Db 220 MKYRDWIIRSKFEWHTLSKEYETQK--VPKENAENYLIKISEKNDKAVSLLNLCDAEY 277

Qy 416 DDLCDCRYTATIIKSPNGPAKNDVDIASQINVNDLRGFGNYKS--NNEKSNWCTGTFT 473
Db 278 SKYCDCKHTTLVKSVLGNDNTIKREHIDLDDDFKFGCKNSVDNTNKTWCKPKYK 337

Qy 474 NKPGTCPPRQTLCLGRVTLHRGHEEDYKEHLLGASIEAQLLYKYKEDENALCS 533
Db 338 LSTKDVCPVPRQELCIGNDRIYDKNLLMIKEHILAIAIESRLKRYKNDKCKVCK 397

Qy 534 IIQNSYADLADIIKGSDDIIKDYKQKMEENLNKVNKDKRNEESLKI PREKWDENKENV 593
Db 398 IINKTFADIRDIIGTGYNDLSNRKLVGKINTNSVHRNKQNDKLFREDWVWVIKKDV 457

Qy 594 WKMSAVLNKNCETKQYDKFKQIPQIRWPKWGGDDCEKKEKIIYSFESKCKKDC 653
Db 458 WNVISWVFKDVTCKE--DDIENIPQFRWSEWGGDDYCDQKTKMI---ETLKVECKBCKPC 513

Db 211 VKNSYLDKDVIFGTDLD-----KNNISKLVESLSKRFK-KDSSVLN--PTAMRRYVGR 263
Qy 290 LWNHMI--VNHKGNISKECAIIPAEPEQJNLWIKEW---NENFLMEKRLFLNAIKDKCVE 344
Db 264 LMKTMIQPYAHLG-----CRKPDENEQJINRWILEWGKYNCRMLMKEKEL--LTGCSV 315
Qy 345 NKKEACFGGCLRLPCSSYTPMKKSXTOMEVLNLYKK-----KNSGVDKNFLNDLF 397
Db 316 NRKSDCSTGCNNECYTYRSLNQRYSVILGKKYIKVRYTIFRKKIYQPNAL--DFL 374
Qy 398 KNNKN---DLDDFFKNE-KEYDDLCDCRYTATIIKSFNLGPAKNDDVDIASOINVNDLR 452
Db 375 KJNCSECKDIDFPFFFEFYKYEKWC-----QSYIDLKIFQNDI- 418
Qy 453 GFQCNYSKNE-----KSWNCTGFTNKF-----PGTCBPPRQTLCLGR 492
Db 419 ---CSFNAQDVTSSDKRFLCKEKEPKWKCD---KNSFETVHHKGVCSVPRQGPCLGN 472
Qy 493 -TYLLHGRHEEDYKEH-----LLGASIEYAQLLYKYKE-KDENALCSLIQNSYADLA 543
Db 473 LNYLL--NDDIYNVHNSQLLIEIMASKQEGKLLWKKHGTILDNQNAKYINDSYVDYK 529
Qy 544 DIIKGSDIIDYKGMKEENL-----KVNKDKRNEESLKI PREKWDENKENVW 594
Db 530 DIVIGNDLWNNISIKVQNNLNLIFERNFGYKVCGRKL--FKTIKELKNVWILNRKQW 587
Qy 595 KVM----SAVLKNETCKVDKQKIPQFLWPKWEGDGFCEKREKIKYISFSPKVECK- 649
Db 588 ESMRCGIDEVDQRKTCERIDELENMFQFPWFQMAHFFC---KEKEYWELKNDKCTG 644
Qy 650 ---KDCDENTCKNCKSEYKWIIDLKSEYKQVDKTKDKNKQMYNDIBVRKKEANY 706
Db 645 NNGSLCQDKTCQNVCTNMNTYTRKLAYEQSVKY--DKDRKLF---SLAKDKNVTTF 699
Qy 707 LKEKSKCKQVNFDDKLFNESPNEVEDMCKCD---EIKYLNIEIKYPTKHDIYDI--- 759
Db 700 LKENAKNSIDF-TKIFDQDLKLFKERCSMDQVLEVRNKKEMLSIDNSSEDATDISEK 758
Qy 760 -----DTFS 763
Db 759 NGEELYVNHNSVSVASGNKEIEKSDEKPEKAQTNGTLTVRTDKDNRNKGKDTAT 818
Qy 764 DT----- 765
Db 819 DTKNSPENLKVQEHGTNGETIKEBPPKLPESSETLSQEQLEABAQKQKEEPPKKQBE 878
Qy 766 -----FGDCTPISINANINEQQSGK 785
Db 879 EPKKQBEEXKQEQKQBEEXKQBEEXKQBEEXKQBEEXKQBEEXKQBEEXKQBEEXKQ 938
Qy 786 D-----TSNTGNETSPVSH-EPESDAAIN--VEKLSGDDESSETRGI 827
Db 939 EQNVKSSSPVVPQETTSNGSSQ--DTKISSTEPNNSVVDRAATDSMNLDPKRVNENM 996
Qy 828 LDIN-----DSVTNNVNEVIDA-----SNTQ-----GSVNTS 856
Db 997 SDPNTNTEPDASLKDCKEVDADAKKELQSTVSRIESNEQVQSTPPDPTTVEKGVGDKA 1056
Qy 857 DI-----TNGHSESSLN-----RTTN----- 872
Db 1057 EMLTSPHATDNSESGNLNPTDDIKTTDGVVKEQBEILGGESATETSKNLEKPKDVEPS 1116
Qy 873 ---AQDIKIGRNEQSD-----NQENSS-----HSSDNGSLTIGQ 906
Db 1117 HEISEPVLSGTTGKESELLKSKSIETKGETDPRSNDQEDATDDVENSRRDNNLS--- 1173
Qy 907 VPSEDNTQNTVDSQP-----HRTPNALASLPDDKDKINEIEGFDSS 948
Db 1174 -NSVDNQSNVLNREDPIATSETVVSEPEDSSRIITTEVPSTTVK-PPDEKRSESVGKEA 1231
Qy 949 RD-----SENGRG-DTTSNTHDVRRTNIVSER-----RVNSHDFIR 983

Db 1232 KEIKVEPVVPRALCEPMENSVSVQSPNVEDVEKETLISENGLHNDTHRCNISEKDLID 1291
Qy 984 NGMANNNAHHQYITQIENNG-IIRGOEESAGN-----SVNYKDNPKRSNFSSE----- 1030
Db 1292 IHLRLNEAGSTILDDRRNGEMTEGSESDVGLQEHNFSTQQKDEKDFDQIADREKEEI 1351
Qy 1031 -----NDH-----KKNIQEYNSRDTKVRREIIR 1054
Db 1352 QKLLNTHGHEDEVDLWMDRTEDSMGVSNSHLYYNLSSEKMEQYNNRDASKREBILN 1411
Qy 1055 LSKQKCNNEYSMEYCTYSDERNSSPGCSRERKKLCCOISDYCLKYFNFSYIEYNCI 1114
Db 1412 RSNWNTCSNEHSUKYQYMERNKDILLETCEKGLHLCCEISDYCLKFPFKSIEYFDCT 1471
Qy 1115 KSEIKSPYKCFKSEGOSSI 1134
Db 1472 QKEFDPTYNCFRKQKQFTSM 1491

RESULT 10
US-10-153-273-6
; Sequence 6, Application US/10153273
; Publication No. US20020169305A1
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; Chitnib, Chetan
; Miller, Louis H.
; Peterson, David S.
; Su, Xin-zhaun
; Wellem, Thomas B.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/153,273
; FILING DATE: 21-May-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/210,288
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Fuller, Michael
; REGISTRATION NUMBER: 36,516
; REFERENCE/DOCKET NUMBER: NIH121.1FWDV1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 749 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-153-273-6

Query Match 11.7%; Score 761.5; DB 4; Length 749;

	Best Local Similarity	27.3%;	Pred. No. 3e-34;	
	Matches	207;	Conservative 148;	Mismatches 282;
			Indels 121;	Gaps 31;
Qy	359	CSYSYTFMKEKSKTOMEVLTWLYKKNGSGVDKNFNFLNDLFFKQNNKNDLDDFFKN--EKEYDD	417	
Db	30	CRSGPEFNK-----IIPERNVQI---HISNIFKEYKENVVDIILPTGLNVEYNN	74	
Qy	418	LC-DCRYTAIIKSLFNGPKAKNDVDIASQINVDNLRGFGCNK--SNNEKSWNC--TGTFTN	474	
Db	75	FCKEKPELVAAKYNLKPAPNAKSPRIYKSEHEBESSVFGCKTKISKVKKWCNYSNNKVT	134	
Qy	475	KFGCTCEPPRQTLCLGRVTLHHRGHEEDYKEHLLGASIEAQLLYKYEKEDENALCSI	534	
Db	135	KPEGVCPPRRQQLCLGYIFLIRDNBEGLDKHINKAANYEAMHLKEKYENAGDGKICNA	199	
Qy	535	IQNSYADLADIIKGSDIIKDYVGKKMEENLNKV-----NKKCKNEESLKFREFKWDEN	589	
Db	195	ILGSYADIGIVRLGDVWRDINTNKLSEKQKIPMGGNSRKKQNDNN--ERNKWWEQ	251	
Qy	590	KENVKVMYSAVLKNKETCKDYDFQKIPQFLRWFKWGGDDFCERKEKIYSEFVKVECK	649	
Db	252	RNLISWSMWXHIPKGTCKGRHNNEKIPQFLRWLKEWGDFCEBEGTEVKLEKI---CE	308	
Qy	650	KKDCDENTCKKCSYKWKIDLKSEYEKQVDKYTKOK--NKK--MYDNDDEVNKKZANV	705	
Db	309	NKQCEKKCKNACSSYBKWKIERKNEYNLQSKFDFSDKGLNKNLNYKFPDSK-----A	363	
Qy	706	YLKESKECKDVNFDDKTIENESPNEYEDMCKCD--EIKYLNRIK--YPKTKHDIYDI	759	
Db	364	YLKSESQCNISIEFNDETFT--FPNKYKBAKVCENPSSSKALKPIKTNVPFIESSKSEL	422	
Qy	760	DTFSDTFDGTPTISINANINEQSGK--DTSNTGNSETSQSPVSHPESDAANV-----	812	
Db	423	SSLTDRKSNTPNSGGNGYGRQIRKDDVHDCPKVK--SGEKEVPKIDAAVTEHEFT	481	
Qy	813	-----EKLGDSESSETRGLINDPS--VTNNVNEVHDASNTQGSVNTSDITNG	861	
Db	482	SNRNDIEGKSKGDHSPVHSKDIKNEEPORVVSLENLPKIEKXWSSDSIPIT-----	535	
Qy	862	HSSESLNRTTNAQ-----IKIGRSGNEQS-----DNQNSHSSDNGSLTIGQVPS	909	
Db	536	HIIEAKGQSSNSDNDPAVVSGRGSKOVNLHTSERIKENESEGVTKTDDSSKSIKISIPS	595	
Qy	910	EDN-----TQN-TYDSQNPHRDTENALASLPDDOK--INEIEGFPSSRDSSENGRDTTNT	962	
Db	596	DQNNHSDLSQANEDSNQGNKETNP-----PSTEKLNKEIHYKTSDDSDHGSKIKSEIEP	651	
Qy	963	HDVVRTNIVISERR-----VNSHDFIRNGMANNNAHQYITQIENNGGIIRGQE	1009	
Db	652	KELTEESPDKKTESAAIGDKNHVESYKSADIFQSEIHNSDNRDRIVSE-----SVV--Q	704	
Qy	1010	ESAGNSVNYK-----DNPGRNSFSENHDKNIQBYNSR	1043	
Db	705	DSGSGSMSTESIRTDN---KDFKTSIEDIAPSINGNSR	739	

RESULT 11

US-10-153-273-2 ; Sequence 2, Application US/10153273
; Publication No. US20020169305A1

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.
Chitnis, Chetan
Miller, Louis H.
Peterson, David S.
Su, Xin-zhaun
Willems, Thomas B.

WEITZMAN, THOMAS B.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

1 TITLE OF INVENTION: BINDING DOMAINS FROM *PLASMODIUM VIVAX*
2
3
4 AND *PLASMODIUM FALCIPARUM* ERYTHROCYTE BINDING PROTEINS
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:

CONTACT ORDERS/SEE ABOVE:
;
;
; ADDRESS: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
;

CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/153,273
Filing DATE: 21-May-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/210,288
Filing DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael
REGISTRATION NUMBER: 36,516
REFERENCE/DOCKET NUMBER: NIH121.1FWDV1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1115 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium vivax
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-153-273-2

Query Match 11.6% Score 754 DB 4: Length 1115;

Best Local Similarity 24.5%; Pred. No. 1.3e+33;
 Query Match 11.6%; Score 734; DB 4; length 1113;
 Matches 296; Conservative 177; Mismatches 416;
 Indels 318; Gaps 57;

Qy		199	PKEIFLISVNTAEKLLYNKGEGKDPSPFCNELRNSFSDFRSSFI GDDMDFGGNTDRVKGY	258
Db	:	22	YKODFSIT-----LNYHEGGKYLIILKH-----KLEKANNRDVCNFF	59
Qy	:	259	IN-TKPSDYKKNVEKLNNIKKEWIB-KNKANLNWHMIVNHKGNISKCAIIPAEPOI	316
Db	:	60	LHFSQVNVLLERTIEFLLECCKNVKGENGYKL-----AKGH-----HC-----VEEDNL	105
Qy	:	317	NLWTIKEWENPLMEKKBLFLNLKDCKVENKGYACFGGCRLPCS-SYTSFWICKSKTOMEV	375
Db	:	106	ERWLQGTNER-----RSEBNIKYK-YGVTELKI KYAQMNKRSSRIILKES	149
Qy	:	376	LJ-----TNLYKKKSQVDKNNFLNDLPKKNKNNDLDDFPKNEKEYDDLCCRY-----	423
Db	:	150	IYGHNFTGGNSYMEGKOGDKTG-----BEKDGEHKTDSTDNKGANNVLMLDIETSSN	204
Qy	:	424	-----TATIIKSFUNGPAKN-----DVD-----IASQINVNDLRGF-----G	455
Db	:	205	GQPAGTLDNVLBFVTGHEGNSRKNSSNGNPYDIDHKKTISSAI-IN-HAPLQNTVMKN	261
Qy	:	456	CNYK-SNNEKSNWCCTGTFTKNFPGTCPPPRQTCLGRTYILLHRCHEDY-----KEH	507
Db	:	262	CNYKRRRBRDWDC-----NTKKDVCPIDRRYQLCMKELTNLVNNTDTNFHRDITFRKLY	316
Qy	:	508	LLGASIIYEA-----QLLKYYKYEKENALCSIIQNSYADLADIIGSDIIDIYYGKK	559
Db	:	317	LKRKLIYDAAVEGDLLLKLANRY-NKD---FCKDIRWSLGDFGDIINGTDMEGIGYSKV	372
Qy	:	560	MEENLNKY-NKDKRNBSLKI FRKXWDENKENVNWKMSAVLNK-----ETCKDYDK	612
Db	:	373	VENNLRIFGDEKAQR-----BKWMNESKAQIWTAMTVSVKRLRGFNFTWICKLNYA	427
Qy	:	613	FOKIPOFLRFPEWGDDPCEKRKEKLYSFEFKYBEC-----KKKODENTCNKCSBY	665

Db 428 VNIIEPQIYRWIREGRDYVSELPEV---QKLKCKDGKYNITDKVKCKVPPCQACKSY 484
Qy 666 KKWIDLKSEYKQVDKYTKDKNKKMVDNIDEVKNKEANV-----YLKESKECKOVNFD 720
Db 485 DOWITREKKNQWVLSNKFISVNAE-----KVQTAGIVTPVDILKQBELDEFNEVAPE 536
Qy 721 DKLFNEPNEVEDMCKKDEIKYLNELIKYPTKHDIYDIDTFSTFG-DGTPI----- 772
Db 537 NEI-NKRDGAYIELC-----VCSVEEAK-KNTQEVWTVNVDNAKSOATNSNPISQPVDS 589
Qy 773 -----SINANINEQSGKDTNTGNSETSD-----SPVSGHE-----PRESDAAINV 812
Db 590 KAEPKPGDSTHGVN-----SGQDSSTGKAVTGCGQNGQTPAESDVORSIDIAESVSAKV 646
Qy 813 E-----KLSGD-----ESSSETRGILDINDPSTVNNVNEVHDASNTQGSV----- 852
Db 647 DPQKSVSKRSDDTASVTGIAEAGKENLGASNSRSESTVEANSFGDDTVNSASIPVVSGE 706
Qy 853 -----SNTSDITNGHSESSLNRTTNAQDIKIGRSG-----NEQSDNQENSHSS 896
Db 707 NPLVTPYNGLRHSKDNSSDGPAESMANPDSNSK-----GETGKGQDNDMAKATKDSNSS 762
Qy 897 DNGSGLT-----IGOVPSEDNTQNTYDSQNPHRDTPNALASLPDDDKINEIEG 944
Db 763 DGTSSATGDTTDAVDREINKGVPEDRDKTVGSKDGGEDNSANKDAATVVGEDIRENSA 822
Qy 945 FDSRD-----SENG-----RGDTTSN---THDVRRTNIV 971
Db 823 GGSNDNRSKNDTEKNGASTPDSKQSDATLSKTESLESTESGDRTTNDTNSLENGKG 882
Qy 972 SERRVNSHDFTRNGMANNAAHQVIT-----QIENGLIIRGOESAGNSVNYKDNPK 1023
Db 883 KENDLQKHDFKSDNTPNEEPSNQDTDAEGHDSIKND-----KAEERKHKMKTFTKNTN 939
Qy 1024 RSNPSSN-----DHKNITQENSRDTRKVRBEIKLSKONKCNNEYSMEYCTYSDERNSSP 1080
Db 940 SHLNSNNLSNGKLDIKEKYRDVKATREDIILMSSVRKKNNISLEYCNVEDKISS- 998
Qy 1081 GPCSRERKKLCCQISDYCLKYFNFYSEIYNYNCISKEIKSPEYKCFKSEG-QSSIPYFAA 1139
Db 999 NTCRSREKSNLCCSISDFCLNYFDVYSVEYLSCKMKEFEDPSYKCFKGGFKIDKTYFAA 1058
Qy 1140 GGLVWVILLSSASRKGKNEEDIGESNIEATFEE-NNYLNKLSRI-----FNQEVQE 1193
Db 1059 AGAL--LILLIASRKMKNDS-----EATFNEFEYCDNIHRIPLMPNNIEHQP 1108
Qy 1194 TNISDYS 1200
Db 1109 STPLDYS 1115

RESULT 12
US-09-924-154-16
; Sequence 16, Application US/09924154
; Patent No. US20020127241A1
; GENERAL INFORMATION:
; APPLICANT: Narum, David L.
; APPLICANT: Sim, Kim L.
; TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use
; FILE REFERENCE: 05213-0465 43170-262105
; CURRENT APPLICATION NUMBER: US/09/924,154
; PRIORITY FILING DATE: 2001-08-07
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Mammalian
US-09-924-154-16

Query Match 11.6%; Score 750.5; DB 3; Length 972;
Best Local Similarity 25.8%; Pred. No. 1.7e-33;
Matches 251; Conservative 143; Mismatches 329; Indels 257; Gaps 44;
Qy 355 CRLPCCSYTFSFMKSKSTQMEVLNLTNLYKK--KNSGVDKQNFN-----DLFKKNKNDD 407
Db 35 CKSECKYKAWIDKKNDFILSEIYLKYNKASLSYKTAPEYILKQKWDYKELNFFSIFD 94
Qy 408 FFKNEKEYDDLCDRCYATATIISFLNPAKNDV-----DIASQINVNDLUG-FGCNYKS 460
Db 95 QL-NAKYNK-CICQ-----NKKIENNALYKIEDICNNTKVKISIVGELYCKEK- 141
Qy 461 NNEKSNWCTGTFTNKPFGTCPTCEPPRQTLCLGRYTLHRGHEEDYK-----EHLG 510
Db 142 GNDKIQCINEHITKDPFVCGPPRQQLCLGNL-----DKDEFKNVNDLKKFLNELILG 195
Qy 511 ASIYEALQLKYKYEK-----DENALCSIIQNSYADLADIIGSDIISKDYKKEEN 563
Db 196 IR-DEGKFLIEKRYKNWHENMYLDERA-CYLYNSFDDYKXIIILGKDMWRDPSIKTENI 253
Qy 564 L-----NKYNKDKKRNEESLKIIFREKMWDEKENVMKMSA-VLNKKT--C--K 608
Db 254 LKGNPEGIKANIVSMYPSYADLSLDEFKHWBQNKQLWEAISCEPYKGNHTGVCLEMD 313
Qy 609 DYDKFOKIQPLAWPKWGDGDFCEKKEKIKYSESPKVEC-----KKKDCDEN-----T 657
Db 314 DNDN-----QYLWPFREKWNDFC---IDKLKWNVDVIKEPCIDKKVSPKSENPSPDVATV 365
Qy 658 CKNKSEYKWIIDLKSEYKQVDKYTKDKNKKMVDNIDEVKNKEANVILKESKECKOV 717
Db 366 CNKSCDYDKWIIINKKEYKQWSSKYRDRS--LFNVN--IQMLKPWEYLSMKCTEC-TC 420
Qy 718 NFDKIFNEPNEYEDMCKKC-----DEIK-----YLNIEIKYPTKHDI----- 756
Db 421 NLDQTQTVYKGYEDICKSTVKPYDPEDIKDEFNEPSLVNPLSLTSQDVTERTVSVD 480
Qy 757 -----YDIDTFSDTFGDTPI 772
Db 481 DVLISIKENVLDKPKPGQTQSSHVDQVGNPRESESKPGSANGREDPSTESSTYNDGV-I 539
Qy 773 SINANINEQSGKDTNT-----CNSETSDSPVSG-HEPESDAA- 809
Db 540 TSSSSILG-SSSGRDVSPPVGVGDEHEAKELLPPQKIIDGVTSQDSESTLSQHKESSEQ 598
Qy 810 --INVEKLSGDESSSTRGILDINDPSTVNNVNEVHDASNTQGSVNTSDITNGHSE-- 864
Db 599 HNLGSSLSRHSNQDEERSI-----ITSQVE--HGTNSLFGSQIQDQETILGESEPLT 649
Qy 865 -SSLNRTTNAQDIKIGRSGNEQSDN--QENSSHSDNSGSLTIGQVPSD-----NT 913
Db 650 TSPPEHETSKMDTHAGGKNMEQVRNASVDSSESMSNGRGLKTKEMKGEVITGITSKND 709
Qy 914 QNTYDSQNPHRDT--PNALASLPDDDKINEIEGF-----DSSRDSDNGSGDT 958
Db 710 INLEDSITVSRQNKIENSGDNTQCKEHIINVQMDXHELENPTTSERCDSVLESEFSLNR 769
Qy 959 TSNTHDVRRTNIVSERRVNSHDFIRNGMANNAAHQVITO-----IENNGI 1004
Db 770 TSHTHDNNRIETTAENNI-----GGLSNVNVHDGRDSQRNRMHINSRHRHGSLESDIV 822
Qy 1005 IRG-----QESAGNSVNYKONPKRSNFSSENHKKNOEYNSRQTKRVRESIILK 1054
Db 823 VRGDDISNTEGGBEEBEDANTLKY-----PRNLNNKNSRTYNTIEEYIYRQVNVKADIMR 878
Qy 1055 LSKQNKCNNEYSMEYCTYSDERNSSPQCSRERKKLCCQISDYCLKYFNFYSEIYNYNCI 1114
Db 879 SYKSNECTNNLSNYCS-KLKEKESLNTCTNEDSKRLCCSISDYCMKFFNFSNGHSCM 937
Qy 1115 KSEIKSPYKCFKSEGQSSI 1134
Db 938 RKEFSNHAYKCFAGKGFSSM 957

QY	452	RGFGC-NYKSNNEKSWN	CTG-----TFTNKPGTCE-----	PP	483
DB	1610	KHCGCTKFTIENDKYTNISSKDKCKGLVKEANTGAIKWQNGPNYNNL	KELTEDVLFP	1669	
QY	484	RQOTLCGRTYLLHRGH-----	BEDYKEHLLGASIEYAQALLKYKYKKEKENAL	531	
DB	1670	RLRLICF-----	HALDGNYYTPEVKDENGRLRKLMEVAATEGYNIGQYYKKEKEKEKIKTSD	1726	
QY	532	-----	CSIIQNSYADLADIIGKSDIIKDYGYKKMBENLNKV-NKD-----	KCRNEE	576
DB	1727	AHKYSYEVPPCSAMKSYFDRLDIILIGDNLED-EKQKTEENLKKLPNRNGT	SVGKSGDS	1785	
QY	577	SL-----	KIPREKWDENKENVWVWSAVLK-----	NKETCKDYKPKQIP	617
DB	1786	TTGNPGSTARKPFWNEKENCKVWAMICGYKGRDDGNGSARSARSD	LLKCGSVPSDDY	1845	
QY	618	-----	QFLRWFKEMGDPCKEKEKIYSPESPKVECKKDCDEN-----	TCKN	660
DB	1846	PMGNRDEBTAYQFLWFAWGEDFC-KHKEK-ELEKLVGACNDYTCGDNE	DKRKKCTD	1902	
QY	661	KCSEYKWKIDLKSEYKQVDKYTKDNKKMYDNIDEVKNKB-ANVYLKESK	CEKDVNF	719	
DB	1903	ACTQYKCPSEWKPYEQEKIKYGENKD-KISEHPVAKDAEDAREYLDQK	L-----	1954	
QY	720	DDKIPNESPNRYEDMCKCDEIKYLNKEIKYPTKHDIYDIDTFSDTFGDT	PISINANTIN	779	
DB	1955	--KICENKSGDCBYKMK-----	DVSTQRLTDCNSQNPASLD	1990	
QY	780	EQSGKDTNTGNSETSDSP--VSHEPESDAAINVKLSGDESSS	ETRGILDINDP-SVT	836	
DB	1991	DEP--KEVBGKNCQVPRPPRVRRETPSPRVSLISKATA--SKKEAKTAP	TPKQPKQVE	2046	
QY	837	NNVNEVHDASNTQGSVNSTDITNGHSESSLNRTTNAQDIKIGRSGNEQSD	NOENSSHSS	896	
DB	2047	NLTTEBQATRRRAAQTRKETS-----	TATTTESD--VGTWYKAILSNKPD	SRGGI	2097
QY	897	DNSGSLTIGQVP-----	SDN-----	TONTYDS-----	919
DB	2098	EGCNPKTYGOYPKWGCIYGVKSKENENGICMPPRKKLCINNITQYLN	MYETENKRDNDIKEA	2157	
QY	920	-----	QNPRDTPNALASLPD-----	DKINEIEGFDSSRDS	951
DB	2158	FIKAAIETQFWLKYIIENPAENELONGTIPDEFKRIWYTYGDKDM	FFGTDISNDK	2217	
QY	952	ENGRGDTTNTDVRRTNTVSERRVNSHD-----	FIRNGHANNNAHQ	994	
DB	2218	K-----	LIIVTNSV--TTILNENKKQDKDBELRKIFWEKNKKFIWEGMI	GLTYH--	2269
QY	995	YITQIENNGIIRQOESAGSNVNYKD-----	NPKRNSFSENDDHKNIQBYNSRDT	KRVRE	1050
DB	2270	-LTD-EN-----	EKEKIRDNYNDMTKLTFSLEBFVKRPFQFLRWFTWEAEFC	NKCRKE	2321
QY	1051	EIIKLS---KQNKCNYSMEYCTYSDERNSSPGCSREERKKLCCQI	SDYCLKYPNFYVS	1107	
DB	2322	OLLKLEAGCKBYECNG-----	SNDGKTQBC---AEACVYQNF--	2356	
QY	1108	IEVYNCIKSEIKSPKCPKSGOSSIPYFAAGILWVILLSSASRMKSE	NEEYDIGE	1167	
DB	2357	-----	IKWKITYEQRQEXF-----	KKDKGKKYKIDPSTE	2387
QY	1168	SNIEATFENNLYN-KLSRI	1186		
DB	2388	RDIEKATCAHEYLNMKL	KEL	2407	

RESULT 14

US-10-153-273-12

; Sequence 12, Application US/10153273

; Publication No. US20020169305A1

; GENERAL INFORMATION:

; APPLICANT: Sim, Kim L.

; Chitnis, Chetan

; Miller, Louis H.

Peterson, David S.
 Su, Xin-zhaun
 Welles, Thomas E.
 TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
 AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
 NUMBER OF SEQUENCES: 37
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Knobbe Martens Olson & Bear
 STREET: 620 Newport Center Drive 16th Floor
 CITY: Newport Beach
 STATE: California
 COUNTRY: US
 ZIP: 92660
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/153,273
 FILING DATE: 21-May-2002
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/210,288
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Fuller, Michael
 REGISTRATION NUMBER: 36,516
 REFERENCE/DOCKET NUMBER: NIH121.1FWDV1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 235-8550
 TELEFAX: (619) 235-0176
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2710 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Plasmodium falciparum
 SEQUENCE DESCRIPTION: SEQ ID NO: 12:
 US-10-153-273-12

Query Match 7.6%; Score 495.5; DB 4; Length 2710;
 Best Local Similarity 19.5%; Pred. No. 1.1e-18;
 Matches 268; Conservative 170; Mismatches 482; Indels 453; Gaps 57;
 112 LSENEGSGTNSNFPANTSEISIGKDNKQYTFIQKTHLFAAGIK-RKSIKWICRENSE 170
 912 INENYSNDRSGSGPCT-----GKGDH-----GGVRWRIGTEWSNIEGKK 953
 171 KIT---VCVPRDKIQLCVANFLNRLTM---EKFKIFILISVNTAKL----- 213
 954 QTSYKAVFLPRRHEMCTSNLENLDVSGVTKNDKASHLLGDVQLAAKTAAEIKRYKD 1013
 214 -----LYNKGEGKPSIFCNELRNSFSDFRSSFTGDDM--DFGNTDRVRGYI----- 259
 1014 QNNIQLTDPIQQDQKQAMCAVRYSFADLGLIIRGRWDDEKSDTDMETRLITVFKNIK 1073
 260 -----NTKPSDYKKNVEKUNNIKKWEKKNKANLNNHMLVNHKGNISKECALIPA 311
 1074 EKHGDIKDNPKYTDESCKPAYK--KLRAADWEANRHQVWRAMKCATKGI---CPGMPV 1128
 312 EE--PQINLWTKWENFLMEKRLFLNIDKQVE--NKKYEAC-----FGGCRLLPCSS 361
 1129 DDYIPQLRWTEWAEWYKRAQSEYDKLKIICADCMKSGDKCTQGDVDCGKCAACDK 1188
 362 YTSFMKSKTKQMEVLTNLY-----KKKNSGVDKN-----NFLNDLFFK--- 398
 1189 YKEIEKWNEQWRKISDKYNLLYLAQAKTTSTNPGRTVLGDDDPDQYQWVDFLTFPIHKASI 1248

399 ----- 398
 1249 AARVLVKRAAGSPTETAAAAPIPTYSTAAGYIHOEIGYGCQEQTOFCEKHGATSTSTT 1308
 399 KNNKNDLDDFFKNKEKYDDLDCRCRYATIKSFLNGPAKNDVDIASQINVD--LRFGF- 455
 1309 KENKE--YTFKQPPPYATACDC-----INRSOTEPPKKEENVESACKIVKILEGKNG 1361
 456 -----CNYKSNNEKSWNCTGTFTNKPFGTCBPPRQTLCLGRTYLLHRHEEDYK--EH 507
 1362 RTTVGECNPK-ESYPDWDCKNIDISHDGAECMPRRQKCL--YYIAHESQTEINIKTDDN 1418
 508 LLGASIVEA---QLLKYKYEKENALCSIIQN-----SYADLADIKGSDI 551
 1419 LKDAFIKTAATAETFLSQYKSKNDSEAKILDGLIPSLQFLRSMYTFGGYRDICLNTDI 1478
 552 IKDYYG-KWMEENLNK-VNKDKKRNEESLKIPREKWDENKENVWKMVMSAVL-----K 602
 1479 SKQNDVAKAKDKIGKFFSKGSKSPGSL--RQEWKTNNGPFIWKMALCALTYVTDTD 1536
 603 NKETCD---YDK-----FQKIPQLRMPKEWGGDDFCRKRKEKIYFESFKV 646
 1537 NKRKIKNDYSYDKVNSQNGNPSLEEFAAKQFLRWMIENGEEFCAERQKK--ENIIKD 1593
 647 ECKK---KDCDE--NTCKNKSEYKKWIDLKSEYEKQVDKYT-----KDKNKKMYD 693
 1594 ACNEINSTQOCNDKHCRCNQCRAQYQYVENKKKFSGGTNNFVLKANVQFQDPEYKGYE 1653
 694 NIDEVKNKEANVYLKEKSKCKDVNFDKIFNESPNE-----YEDCKCKDEIKYL 744
 1654 YKGVQPIQGNVYLLOKCDNNKSCMDGNVLSVPKPKPGKYAHKYPEK-DCYQKQKHV 1712
 745 NEIKYPK-----TKHDIYDIDTFSDT-----FGDGTPIISANINAEQ 781
 1713 PSIPPPPPVQPOPEAPTVDVCSIVKTLFKDTNNFSDACGLKYGTAFSSWKCIPSDT 1772
 782 QSGKDTSTNGSETSDSPV-----SHEPESD----- 807
 1773 KSGAG-ATTGKSGSGSGSICIPRRRLRYVVKLQEWATALPOGEGAAPSHRADDLRNF 1831
 808 ---AAINV-----EKLSDGESSESTRGIL-----DINPSPVTNNVNEVHDASNTQ 849
 1832 IQSAAIETFLMDRYKEEKKPQDGGQOALSQLTSTYSDEEDPP-----DKLLQN 1882
 850 GSV-----SNTSDI--TNHSESSLARTTNAQDIKIGRSNEQ 885
 1883 GKIPDPFLRLMFYTLGDIYDLVHGNTSDSGNTNG-----SNNNNIVLEASNGKE 1933
 886 SDN--QENSSHSDNSGSLTI-----GOVPSE--DNTQN-----TYDSQNPHRDT 926
 1934 DMOKIQEKIEQILPKNGGTPLVPKSSAQTPDKWNEHAESIWKGMICALITYTEKNPDTSA 1993
 927 PNALASIPSDDKINEIEGFDSSRSDSENGRDTTSTNHDVRRITNIVSERRVNSHDFIRNGM 986
 1994 RGDENKIEKDEEV--YEKFFGSTADKHGTASTTGT----- 2027
 987 ANNAHQYITQIENNGIIRGOEESAGNSVNYKDNPKRSNFSSENDHKKNIQEYNSRDTK 1046
 2028 -----YKTQDYEVKLEDTSGAKTPSASDTPLLSDFLVRPPYFRYLEEONGQNFCK 2079
 1047 RVREELIKLSKQNK-----CNNEYSMEYCTYSD-ERNSSSPGP 1082
 2080 KRKHKLQAQIKHEKVEENGSGSRGGITROYSGDEACNEMLPKNDGTVPDLEKPSCAKP 2139
 1083 CSR-----BERKKLCQISDYCLKYFNFYFYSIYNNCIKSIKISPEY 1123
 2140 CSSRWKIESKGEFEKQKAEQKQDKCVNGSKHNGFCETILTSSKAKDF 2192

RESULT 15
 US-10-732-923-3351
 ; Sequence 3351, Application US/10732923
 ; Publication No. US20050108791A1

GENERAL INFORMATION:

APPLICANT: Edgerton, Michael D
 TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
 FILE REFERENCE: 38-15(52796)C
 CURRENT APPLICATION NUMBER: US/10/732,923
 CURRENT FILING DATE: 2003-12-10
 PRIOR APPLICATION NUMBER: 2003-12-10
 PRIOR FILING DATE: 2002-12-04
 NUMBER OF SEQ ID NOS: 24149
 SEQ ID NO 3351
 LENGTH: 1985
 TYPE: PRT
 ORGANISM: Plasmodium yoelii yoelii
 US-10-732-923-3351

Query Match 6.3%; Score 411; DB 5; Length 1985;
 Best Local Similarity 19.2%; Pred. No. 4.2e-14;
 Matches 317; Conservative 238; Mismatches 479; Indels 618; Gaps 79;

38 SSDISRVNSPE-----LNNHKTNIYDS-----DYEDVNNKLN 71
 327 NSEIKVNNKNTDNNLNNNNNNYNTYQNRNESKVEHDLGKLRNFOLDNRKHLN 386
 72 SF---VENKSVKRSLSFINNKTYSYDIPPSYSYNDK---FNSLSEHEDNSGNTSN- 125
 387 LFNNKMNESYTNKMDHMDKKNENY---NFSFINEKNIITKLTNDDALNFDKFN 441
 126 -----NFANTSISIGKONKQYTFQKTHLEA-- 153
 442 RKQMDVKSIVCVQASQMARLQHLHIESRQMLETEKMLKKKKENP---EKVGMADK 498
 154 -----CGIKRSIKWICRENSSEKIVCPDRIQICLVANFLNLSRLTMEKEIFLI 205
 499 EKEIDKTHSQIKKELS-INKNNE---IDEKEYVNSIKSYDNAQKELLDKMECILI 554
 206 SVNTEAKL-----YNNN---EKGPSI-----FCNELNSFSDFRSF 241
 555 ENKCKSLVDEYDFGQFNKKIIEBEKEIEQEKKEKKNENLNENRREIDEEKLN 614
 242 IGD-----DMDFGNTDRVKGYINTKESDYKE---KN----- 271
 615 MKQKLEMLKKELESLEKKEKKIIDCEYNNLQKSEELRRNENNLKIKELNRIDKY 674
 272 ---VEKLNNIKKEWKNKANLNMH-----IVNHKGNISKECAIIPASEPQINLIKE 322
 675 NELIDELNKKCB-IENDKMKLNDIQDERIKLLNETNNIKKE-----NEKEINYKEE 727
 323 WNE-----NPLMEKKRLFN-----IKDKVENKKY-EACFGGRLPCSSYTFM 366
 728 IKKERISMINDEVKQKRLMLEDIENTKNIMLEDMEKENTKIEBIENDKRMWIKNIEDEK 787
 367 KSKTOME--VLTNLYKKCN-----SGVDKNNFL-----NDLKKNNKNDLDDPFKKEKD 416
 788 EKYTYLEKFPNLEKESLQKCYDDENNRLOAEINNEKCKINK-ERDNLKQKQVYE 846
 417 DLCDCRYTATIKSFLGPAKNDVDIASQINV-----NDLRGFGCNYS 460
 847 D-----EFRNKCEYEDIRKYNMLEENNNMYRIMKQEBLE-----NYKK 890
 461 N-----NEKSWNCTGFTTNKPGTCPPRRQTLCLGRTYLLHRRHGBEDYKEHLLGASIY 514
 891 NVYLDIEBKD-----KLIVQOEKINLEKENLIVEKEQI 924
 515 EAQLLKTK-YEKDENALCSIIQNSYADLADIIGSDIIDKYKQKKEENLKV----- 567
 925 DIELKNFNFKKEENDIKIRIINLSQKEDLANKEKNI-----EKEKDLKLEIKYDLDA 979
 568 -----NFKKRNESLAKIRKQWENKENVWMSAVLKNKETCKDYKFKQIKPQLRW 622
 980 REGLANDKIQMEKSLFDEQ-----LEKIKQKBEELLYDRELKTKEM--- 1024
 623 FKWGGDFCE-----KRKEIYSFES-FKVECKKCKDCDENTCKNKSEYKK- 667

1025 -----DLIEKGTEIRKKNELNKKKEKLDLSDNELKSYSSKLQDREKKLKEKTKELQKV 1078
 668 ---WIDLKSEYKQV-----DKYT-----KDKNKQY 692
 1079 KDQVDYKSLKEKIQFQMIEREKELDEQTVIQIDRNSLEAKKQFLLIKEKHKDT 1138
 693 DNIDE-----VKNEANVYLKESKEC----- 714
 1139 EYIQEQKLLHQLKKNKDKSLKEKENEINLLKQDCRSKNGKGAIIKSNYDNRLKNN 1198
 715 -----KQVNF-----DKIFESPNEYE 732
 1199 LSRISVSKRVKAKAIGKNKFDIIRTRRSKIYNNYNNLEINNSIQYIDSVINDFKNKP 1258
 733 DMCKKDEIKYLN-----IKYPTKHDIDYDITFSDTFDGTGPISINAMINE 780
 1259 FLKNKNNLSLENFQEQGNMFKYPN-----DYDNSWYDKN-----NRNIN 1306
 781 QSGKQDTSNTGSETSDP-----VSHE-----PESDAAINVEK 814
 1307 NSMKNIITDYGNVYENDEPERENRQNIISNENDYENKKNKDNKKKYPKPPRA-NLEK 1365
 815 -----LSGDESSSETRGILDINDPSVTN-VNVEVHDASNTQG---SVNTSD----- 857
 1366 DSYNVANGNKSERNDFNDYNDNTTNYKKKEDYDNNYGDHMYNNNDNHNNTKGN 1425
 858 -----ITNGHSESLNRTTNAQDIKIGRSG-NE--QSDN-----QENSSHSDNS--GSL 902
 1426 VYENNNNGYNNYNNINNNNTWIKMGSLDLDITKSGRLNLFQENKQOENNMJGSL 1485
 903 TIQGVPSD--NTQNTY-----DSQNPDRTPNALASLPS-----DDKINETEGFDSS 948
 1486 NVANLIESDIIDTKSNVGYHPDNHDKYGENESNRNIMCINKNINFDNDVNNINNNISN 1545
 949 RDS-----ENGRGDTTNTHDVRTNIVSERRVNSHDFIRNGMANNNAHQ-YI 996
 1546 INSKLYMNSVYEEEN---DSFVGEHNNKVRNKHSCQNNN-----ENGMSNKNPEKPSL 1597
 997 TQIENNGIIRGQESAGNSVNYKDNPKRNSFSSNDHKNIQEVNS-RDTKRVREEIILK 1055
 1598 NRUDS--LVNNQMSVQISLNMOSKCANNTWKNLSNKNYSUKOTKR----- 1647
 1056 SKQKCNKNSMEYCYTSDERNSSPGPCSEER-KKLCCQISDYCLKYFNFYSIE----- 1109
 1648 -----NNND-----NYDDNNESP---KKEDNFELEELNDG--KSYNSNFIKGD 1691
 1110 -YTNCKS-----ETKSPYKCFKSEGOSSIPYPAAGI----- 1142
 1692 FYTNMEKSLDKSMNETYDHNKNNYNTSNKNMDRKVHTLKNNTNSTNFTVNTNHN 1751
 1143 -----LVVIVLLSSASRMGKSNEEYDICESNIEATFEENNYL-- 1180
 1752 DNTKINHVNRNNGETRKNINLSLMPGNNGKIGMDNNGYN-SDSNI-----MRNQNYLTN 1807
 1181 -NKLRIFNQ8-VQETNIDSYSEYNNKMY 1210
 1808 ENYNSRMNSHAIEQDYNKQKRGKNNENDY 1839

RESULT 16

US-10-732-923-15035
 ; Sequence 15035, Application US/10732923
 ; Publication No. US20050108791A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Edgerton, Michael D
 ; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
 ; FILE REFERENCE: 38-15(52796)C
 ; CURRENT APPLICATION NUMBER: US/10/732,923
 ; CURRENT FILING DATE: 2003-12-10
 ; PRIOR APPLICATION NUMBER: 10/310,154
 ; PRIOR FILING DATE: 2002-12-04
 ; NUMBER OF SEQ ID NOS: 24149
 ; SEQ ID NO 15035

Db 364 MDNEFNPNINSHINRSTTVRSVFPQLEYNVENNEIRKFTNTIGNYMGRENDAI 423
Qy 207 VNTAKLYNNEGKDPISFICNELNSFSD-----PRSSFIGDDMDFGGNTD-----RV 255
Db 424 LNI SPVLEKINRRNRTHOLF--NESSNISTDKLTHKVTSPFLFONINDSNVSNIEPE 482
Qy 256 KGYINTKFSYVKKVKNKLANIKK-----EWKKNKANLWNHMI VNHKGNISKECA--II 309
Db 483 KGGKMEKVEKMEKMEKMEKMEKMEKMEKMEKMEKMEKMEKMEKMEKMEKMEKMEK 542
Qy 310 PAEPQINLWIKENENFMEKRLFLNKKVCEN-----KYYE 349
Db 543 AMDENPIL-----NDSNTINKNIIDAESDKLVQNLDMIKIKLQDDKNKNIDTIKNE 597
Qy 350 ACFGGRLPSCSYTFMKSKTQMEVLTN---LYKKNSGV--DKNPLND-----LPKN 400
Db 598 S-----KRITSLLEKTGSMELCNKI COLKKNVEDVEDKNDTIRBETHQINIDKD 648
Qy 401 NKNDLDDF-----PKNEKEYDDLCDCRYTATIISFLNGPAKNDVDIASQINVNDLRGFG 455
Db 649 KENEINLLYVINIKKKIKKELEAFTNSFVLKX---KMNDEII----- 692
Qy 456 CNYKSNKSNWCTGTFTNKPFGTCPPRQTLCLGRTYLLHRGHEEDYKHEHLLGASIE 515
Db 693 -NIYKNRE--NVFLTITNKL-----EKKR-----LKYAEFKKEKQYKELIM----- 732
Qy 516 AQLLYKYKEDENALCSIIQ-NSVADLADIIGSDIIKDYK-----KMEENLKV 567
Db 733 -----KKEB--LLKIXQINNKKNTFTFNOEQMLKDEIQIKTELKREKIELKQDF 781
Qy 568 NKDKRNEESLKIPREKWDENKENVKMGSAVLKN-----KETCKDY-----DKFKI 616
Db 782 EKTQKINMENLEMEKESFINNLENIENKMEKLEEKYETOIKETEMKYKQIKKEIEKT 841
Qy 617 PQ-----FLRWFKWGGDDFCERK-----EKIYFESFKVECKKQCDENTCKNKC 662
Db 842 KQNAEQNFNSFKYKENL-EKNKNDPFINLLIEKNNEIESFKNDIEQK-----KF 891
Qy 663 SEYKKWID-----LKSEYKQVDKYTKDNKKMYDNIDE-----VKNKEANYLKEKSEK 714
Db 892 KEMKFKENKLLQNNFENKNHPIEQNHII-ENIKKYEYELIKNNEIB-YLKEEMKN 949
Qy 715 K---DVNFDDKIFNESPNEYEDMCKDEIKYLNKIKYPTKTHDIYD-----IDTFSDT 765
Db 950 KIQETENVELKAEKNKHIDMKKELENI-YNVEINLKEKINAEQTLAQMLETQNE 1008
Qy 766 FGDGTPISANINEQQ-----SGKDTSTNGSETSDSPVSHPESDAIAIN 811
Db 1009 YEQ--KLBIQRNEVEQLEIRNEMENKLIBQRKESYDLGSKSEIQLITDKIEKLQWE 1066
Qy 812 VEKLSGDSSESSTRGILDINDPSVTNNVNEVHDASNTQGSVNTSDITNGH-----SESSL 867
Db 1067 IEIINKKIDLE-HNISLIND-----ERENILNQNKNLEDNLKNNEIYNNKINILNDEKL 1121
Qy 868 NRTTNAODIKIGRSNEQSD-----NOENS 892
Db 1122 KLEKEIEHK--ENGAKESDQIREQADLLOEBEIDRIKKESEKKNYIYQVNEINEEYE 1179
Qy 893 SHSDNSGSLTIGQVPS-----BDNTQNTYDSQNPHRDTPNALASLPSDDKINEI-EG 944
Db 1180 EKKKEYNDLLEKANVSNKQLEKCEENQKINEVEDMIKMLENQEVLVT-KKIQELNED 1238
Qy 945 FDSRDS--ENRGDITSN--THDVRTNIVSERVNSHDFIRNGMANNAH-----HOYI 996
Db 1239 FLKKKEAFDNKNDLLKNYEHAIENKHIKEQLEN-----FTNSNEEKISQIKNQYE 1290
Qy 997 TOI-----ENN-----GIIRGOESAG-----NSVYKDNPKR----- 1024
Db 1291 TOIKEMQIHPELLNETKNKNEMLIEDIVKQNDKEQSIQHMNSVNTSLVTKNQLNLNI 1350
Qy 1025 -----SNFS-----SENDHKKNIOEYNSRD-----TKRVRE----- 1050

Db 1351 TEIKNELSNFOEKYDKLVKKNHNLKNBEQTKFLDKILDENNVNKLMEVTEKEREINIKN 1410
Qy 1051 -EIIKLKSKONKNNEYSMEYCTYSDE-----RNSSPCPCSR 1086
Db 1411 VNIILK-SLQOINEQTKL-YKEYTDELKDEILQKKQKNSIDSNVPIDANSSINNALLIE 1468
Qy 1087 ERKKLCCQ-----ISDYCLKYFNFYSGIYYNCKISRIKSP 1123
Db 1469 ENKKLKLQNEVISTKNDTMSVLDNLTKRLSF-----IESKIKBEKY 1510

RESULT 18
US-10-087-013-8
; Sequence 8, Application US/10087013
; Publication No. US20040062769A1
; GENERAL INFORMATION:
; APPLICANT: Arthur Scherf
; APPLICANT: Louis H. Miller
; APPLICANT: Benoit Gamain
; APPLICANT: Dror I. Baruch
; APPLICANT: Pierre Buffet
; APPLICANT: Christine Scheidig
; APPLICANT: Jurg Gysin
; APPLICANT: Bruno Pouvelle
; APPLICANT: No. US20040062769A1utaka Fujii
; APPLICANT: Joseph Smith
; TITLE OF INVENTION: IDENTIFICATION OF THE DOMAIN OF
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ERYTHROCYTE MEMBRANE PROTEIN 1
; TITLE OF INVENTION: (PFEMP1) THAT MEDIATES ADHESION TO CHONDROITIN SULFATE A
; FILE REFERENCE: NIH176.001C1
; CURRENT APPLICATION NUMBER: US/10/087,013
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: PCT/US00/24195
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/152,023
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-087-013-8

Query Match 5.6%; Score 363; DB 4; Length 407;
Best Local Similarity 28.3%; Pred. No. 2.9e-12;
Matches 110; Conservative 49; Mismatches 124; Indels 106; Gaps 19;

Qy 456 CNYK---SNNEKSNWCTGTFTNKPFGTCPPRQTLCL-GRTYLLHRGHEEDYKHEHLLG 510
Db 12 CNTKYPTKNDYPGWNTCTDKVINREEGSCMPPRQKLCIHNLHSEKATETELKAKAFIE 71
Qy 511 ASIYEAQLLYKYKE--KQENAL-----CSIIQ-----SYADLA 543
Db 72 CAATETFLWDKYKDEKKDEKTEGGGTSDDPDPQKKLEGGTIPEDPKRMFTYGYDR 131
Qy 544 DIIIGSDIIDY-----YKKMEENLKNVKNKKNESLKIIFREKWDENKENVKQMS 598
Db 132 DFLFTDISGHGKESALGKKI-DLSPK-NGDQKSPSGKTP---TEWWDYGPDIWKGMV 186
Qy 599 AVLK-----NKECK-----DYDKFQKI-----POFLRWFKWGGDDFCBKRKEKI 638
Db 187 CGLSHHINKNGKEQLKRLNTDNKYTKISSKLEDFASRPQFLRWFIWGDQFC---REVR 243
Qy 639 YSFSFKVECKKDC--DEN-----TCNKKSEYKWKIDLKKSEYKQVDKYTKDNKKQMY 692
Db 244 VKINQLATGCVNEYCGSQENGKEACKACNACEYKSLKDWKQYEQQTAKFDKDKCKKF 303
Qy 693 D-----NIDEVKNKEANYLKEKSK-----ECKQVNFDDKIFNES 727
Db 304 DGTSAFDVAASVSHVYLOELKNLCTKGDCACMEKPSAQDBETELLGNGYFFPEAMDYP 363
Qy 728 PNEYEDMCK-----KCDB--IKYLN 746

```
Db 364 PKEIGERCKCAIPSEPMSCVEQIAKHLRE 392 ||| : ||| ||| : ||| ||| : |||
; Sequence 8, Application US/10153273
; Publication No. US20020169305A1
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; Chitnis, Chetan
; Miller, Louis H.
; Peterson, David S.
; Su, Xin-zhaun
; Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/153,273
; FILING DATE: 21-May-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/210,288
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Fuller, Michael
; REGISTRATION NUMBER: 36,516
; REFERENCE/DOCKET NUMBER: NIH121.1FWDV1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 921 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-153-273-8
Query Match 5.6%; Score 360.5; DB 4; Length 921;
Best Local Similarity 22.4%; Pred. No. 1.1e-11;
Matches 166; Conservative 83; Mismatches 231; Indels 261; Gaps 33;
Qy 217 KN-EGKDPISFCNELRNSPDSFGDDMDPGNTDRVKYINTKFSYYKEKN----- 271
Db ||| : ||| ||| : ||| ||| : ||| ||| : |||
58 KNCNGGDTGKGCNG-SNGKNGNDYITASDIENGNS---IGNIDMVVSD--KDANGFNG 111
Qy 272 -----VEKLNINIKKEWKEKNK-ANLWNHMTVNHGNIKSECAIIPAEEO---INLWIK 321
Db ||| : ||| ||| : ||| ||| : ||| ||| : |||
112 LDACGSANIPKGIKEQWKCAKVGCLDVCGLKNKGSGIDK-----DQKIIIRALLK 164
Qy 322 EWNEFMLEKRLFLNITKDCVENKKYEAFCGGRLPCSSYTSFMKSKTQMEVLTNLYK 381
Db ||| : ||| ||| : ||| ||| : ||| ||| : |||
165 RVEYFLEDYNNKINAKI-SHCTKKDNSTCTNDCPNKCTCVEEINQRTWEKNKKHYK 223
;
382 KKNSGVDKN--NFLDLF-----KQNK-----NDLDDFFKNEKEYD 416
; ||| : ||| ||| : ||| ||| : ||| ||| : |||
224 TQENGENNMKSLVTDILGALQPSDVNKAIKPCSLGTAFESFGLGADNSKKEGEDY 283
; ||| : ||| ||| : ||| ||| : ||| ||| : |||
417 DL-----CDCRYTATIISKFLNG-----NGGKSCPTPLDNTLEBEPLEENQVEAPNI 434
; ||| : ||| ||| : ||| ||| : ||| ||| : |||
284 DLVLCMLKNLEKIQECKKHGTSVE---NGGKSCPTPLDNTLEBEPLEENQVEAPNI 340
; ||| : ||| ||| : ||| ||| : ||| ||| : |||
435 -----PA-----KNDVDIAS 444
; ||| : ||| ||| : ||| ||| : ||| ||| : |||
341 CPKQTVEDKKKEEETCTPASVPPEKVPVHVARWRTFTTPEVFKIWRGRNKTTCIVA 400
; ||| : ||| ||| : ||| ||| : ||| ||| : |||
445 QINVDLRG---PGCNKSNNEKSNWCTGTFTNKP---GTCEPPRROTICL----- 490
; ||| : ||| ||| : ||| ||| : ||| ||| : |||
401 EM-LKDKNGRRTTVGECYRKETYSEWTC---ESKIMGQHGACIPPRRQKLCILHLEKIM 456
; ||| : ||| ||| : ||| ||| : ||| ||| : |||
491 -----GRTYLL-----HRGHEEDYKEHLLGASIVYEAQLLKVKYKEKDE 528
; ||| : ||| ||| : ||| ||| : ||| ||| : |||
457 TWTNELKAFIKCAAAETFLWQYKDKNGNAEDLDEKLGKGIIPDFKQMPY----- 511
; ||| : ||| ||| : ||| ||| : ||| ||| : |||
529 NALCSIIQNSYADLADIIGSDIHKDYGKMEENLKV--NKDKKNEESLKI-FREKW 585
; ||| : ||| ||| : ||| ||| : ||| ||| : |||
512 -----TFADYRDICLGTDISSK---KDTSGVGKVKCNIDDDVFKISNSIRYKSW 559
; ||| : ||| ||| : ||| ||| : ||| ||| : |||
586 WDEKENVWVKMSAVLK-----NKETCK-----DYDKFQKIP 617
; ||| : ||| ||| : ||| ||| : ||| ||| : |||
560 WETNGPVIWEGMLCALSIDTSLANNVAPETHKLTGEGNNNFKVIFGSDSSTTLTKFSERP 619
; ||| : ||| ||| : ||| ||| : ||| ||| : |||
618 QFLRWFKWGDDDFCEKRKEKIYSFESFKVECKK-KDCDEN-----TCNKKCSEYK 666
; ||| : ||| ||| : ||| ||| : ||| ||| : |||
620 QFLRWLTWGENFCKEQKKE-----YKULLACKDCDVGDKGKNGKCVACKDQCKQYH 673
; ||| : ||| ||| : ||| ||| : ||| ||| : |||
667 KWIDLKKSBEYKQVDKYTKDKNKKMYDNIDEVKNK-EANYLKEKSKCK-----DVNF 719
; ||| : ||| ||| : ||| ||| : ||| ||| : |||
674 SWIGIWIDNYKKQGRYTEVKKIPLYKEDKDKVNSDDARDYLTQLQNMKCVNGTTDENC 733
; ||| : ||| ||| : ||| ||| : ||| ||| : |||
720 DDKIFNESPEYEDMCKKDE 740
; ||| : ||| ||| : ||| ||| : ||| ||| : |||
734 EYKCMHKSTSTNSDMPESLDE 754
;
RESULT 20
US-10-732-923-22586
; Sequence 22586, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 22586
; LENGTH: 4226
; TYPE: PRT
; ORGANISM: Plasmodium falciparum 3D7
US-10-732-923-22586
Query Match 5.4%; Score 353; DB 5; Length 4226;
Best Local Similarity 18.5%; Pred. No. 1.9e-10;
Matches 259; Conservative 228; Mismatches 452; Indels 460; Gaps 64;
Qy 25 SIIGRTLYNRQDESSDISRVNSPELNNHKNITNYDSYEDVNNKLNISFVENKSVKKRS 84
; ||| : ||| ||| : ||| ||| : ||| ||| : |||
772 SIAGTYGSKCKKKKI--YNNPKSNNNKLNSYD-----KNIFKFKFSSEASIKK--- 820
; ||| : ||| ||| : ||| ||| : ||| ||| : |||
85 LSFI-----NNKTKSYDI-----IPPSYSYRNDKFNLSLSENGNSGNTNS 124
; ||| : ||| ||| : ||| ||| : ||| ||| : |||
821 ISFVRDPSNNVSSKNDLTLDDPTLISDEGNLYLRDKYEHTSDKNDTNKRDGANNNSN 880
; ||| : ||| ||| : ||| ||| : ||| ||| : |||
125 NNPANTSEISIGKDNQYTFIQKRTLH---FACGI-----KRSIKW-----ICRE 167
; ||| : ||| ||| : ||| ||| : ||| ||| : |||
```



```

; APPLICANT: Louis H. Miller
; APPLICANT: Benoit Gamain
; APPLICANT: Dror I. Baruch
; APPLICANT: Pierre Buffet
; APPLICANT: Christine Scheidig
; APPLICANT: Jurg Gysin
; APPLICANT: Bruno Pouvelle
; APPLICANT: No. US2004062769A/utaka Fujii
; APPLICANT: Joseph Smith
; TITLE OF INVENTION: IDENTIFICATION OF THE DOMAIN OF
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ERYTHROCYTE MEMBRANE PROTEIN 1
; TITLE OF INVENTION: (PFMP1) THAT MEDIATES ADHESION TO CHONDROITIN SULFATE A
; FILE REFERENCE: NIH176.001C1
; CURRENT APPLICATION NUMBER: US/10/087,013
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: PCT/US00/24195
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/152,023
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
; US-10-087-013-7

Query Match 5.4%; Score 348; DB 4; Length 294;
Best Local Similarity 32.9%; Pred. No. 1.4e-11;
Matches 94; Conservative 37; Mismatches 87; Indels 68; Gaps 15;

Qy 456 CNYKSNNEKSWNC-TGFTFNKFPOTCEPPRRQTICLGRTYLLHGHEDYK-----EHL 508
Db 13 CNAKRRKKNWQCDKNTFVDGNEGVCMPPRRKRSICIHNLTL-----EQTKNKYQLREAF 67
Qy 509 LGASIIYEAQLLYKYKEDENALCSIIQ-----SYADLADIIGSDIILKDY 555
Db 68 IKCAAKETNLWDYK-NDKNAEBELUKGKIPDFWMRIMFYTPGDFRDFCLENDMGKDV 126
Qy 556 YGKMEENLNKV-NKDKRNBESLKIIFREKWMNDENKENVKMSAVL-----KNKE 605
Db 127 --DKVKNINKVFNSSKRGFK--KIDPENWNNENGPIWNGMLCALIHADTKDSIKND 182
Qy 606 TKRDYDK-----FQIPQPLRFKWDGDDFCBKRKEKISYSPFKVEC 648
Db 183 NYK-YEYVTLAKRDGNGMTLSFAPKPKPLRFVWEYDDYCKEROKYLTVEAS---TC 238
Qy 649 KKKD-----CDENTCKNCKSEYKKWIDLKSEYKQVDKYTKDNK 689
Db 239 KSIDGGQLKDRG-CNNKCBYKMYMRKKKEEWNLQ-DKYIKQRE 282

RESULT 25
US-10-732-923-8668
; Sequence 8668, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 8668
; LENGTH: 2719
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
; US-10-732-923-8668

Query Match 5.2%; Score 338; DB 5; Length 2719;
Best Local Similarity 19.8%; Pred. No. 7.5e-10;

Matches 305; Conservative 199; Mismatches 490; Indels 548; Gaps 70;
Qy 33 NRODESSDISRVNSPELNNHKNIIYDSY-----EDVNNKLINSFVENKSVKKRSLSFI 88
Db 1303 NKKDDGDDILVSKNDEDDDDDDNSVDKYNNNNDNDAIDIDEVMMDEAELESNNNT 1362
Qy 89 NNKTSYDIIPPSYSYRNKFPNSLSENEEDNSGNTNNTNNTFANTSEISIGKNKO-----141
Db 1363 NNNNNNDGTGGSGEMINGQFSS-SNNQNNNNNNNN-----DISLWRSVRKKPQIL 1416
Qy 142 YT-----FIQKRT-----HLFAC-----154
Db 1417 YTPBYGRQEVKQIKHEVRCFICNNYKNHKAIVKDRNGNDIDIDYGDMLNCFRCPTYH 1476
Qy 155 ----GIRKSIK--WIC-----RENSE--KITV-----CVPORKIQLCV 185
Db 1477 KLCEGIKDNVKTWTCWSHCECLCFKSSQCCNLLIHCACTPSCFYCNCFPPDYVRYV 1536
Qy 186 AN-----FLNSRLETMEKFEIFLISVNTAEAKLLYNKNEGKDPDSI 225
Db 1537 GEYYHNLRRQGVNFTPNWVCFKSCKAVEBQK-----RRKMTKEEREHEKQ 1586
Qy 226 FCNELRNSPSDF-----RSSFI-----GDDMDFGENTDRVKGYNLTK 262
Db 1587 LQKELRSQHLHDSQBELEAKKRAQOQLERDKFIENRKRIDALDOQYEDQLRAYENVF 1646
Qy 263 FSDYKKEKVEKLNLIKEMWEKNKANLWNHMTVNHKGNISKECAIIPABEQINLWIK 322
Db 1647 PNNFVK-----ELVNRIEHAKMLKQKIVEDN-----NNNNNNKNTN-----KK 1686
Qy 323 WNFENFLEK---KRLFL--NIKDKVENKYEACFGCRLP-----CSSYTSFMKSK 370
Db 1687 KSTTFLHTKLPSKLLVLCENCKLPCHANYKYP--GKCCYPPPELDKSYNNTSFSQMR 1743
Qy 371 TOME-----VLTNLYKKKNSGVDKNPLNDLFLKKNKNDLD-DPFKNEKEYDLC 419
Db 1744 DGEKRVSEKNDSCSHHFQDKVGEDSNM-----KNITVIGNKIMNSSPHKG 1797
Qy 420 DCRYTATIISKFLNGPAKNDVDIASQINVDL-RGFCGNYKSNNEKSWNC-----TGTF 473
Db 1798 KCEPNEGSTTPSRNDGKKDLMTVTSTNNNNNNNNNNNNNNNNNNNNNNNNNNNNSS 1857
Qy 474 NKFPCTGPPRRQTICLGRTYLLHGHEDY-----KEHLL-----GASIIYEAQLLYKYK 525
Db 1858 NK-----LMSNKGIEYNKDDDEDYNNNESKDKIMKGVGTFLSDGSMRFP-KD 1904
Qy 526 KDNAL-----CSIIQNSYADLADI 546
Db 1905 SPENNANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1962
Qy 547 KGSII-----KDYGK--KMEENLNKVKDKRNBESLKIIFREKWMNDENKENVKMSA 599
Db 1963 KHCDKLTEEBEKEYDEKREKLKELINLLNKKIK-----EETEBEYKNMS- 2008
Qy 600 VLKNETCKDY-DKFOKIQPLRFKWDGDDFCBKR-----EKIYSFESFKVEC 648
Db 2009 VPKFSLYSNYQKADDDILEDCIRSMKWDLIADKKKIIKKEKEKENKKIYS-----2060
Qy 649 KKKDCDENTCKNCKSEYKKWI---DLK-----KSEYKQVD-----KYTK 685
Db 2061 -KQDDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2119
Qy 686 DKNNKMYNDIDEVKNKEANYLKEKSKEDKVDNFDKIFN-----ESPNEYED 733
Db 2120 DMKEKINDELCTISGKRKKKIVKQIKKVKKKGKNDNNDNNNNNNNNNNNNNNNN 2179
Qy 734 MCKKCD--EIKYLNKYP---KTKHDIYDIDTFSDTFGDC-----TPISINAN 777
Db 2180 SSKKDDKOLKLNLLINPPHEKLIKDTNANFIRDSFNEDFLNKAKEFLRVTKSNLIN 2239
Qy 778 INEQSQGKOTSGNSETSDSPVSHPEPESDAAINVEKLSGDESSETRGILDINDPSVTN 837
Db 2240 GNNNNNNKKSNSK-----NINKNS-----SNTK 2266
```


Db 1419 -----ENDNNNMENKGNAN-----STNNKINKINH-----NFQ 1452
Qy 1076 RNSPGPCSRERKLCQISDY-----CLKYFNFYSEYNYNCIKSEIKSPYKCP-- 1126
Db 1453 NNIKTLSBKKEKMYLTIDSINNSFNILDKVNTFHSNYHIALPFLCKYTSYNFLH 1512
Qy 1127 ---KSEG-QSIPYFAGGTLV-----VIVLLSSASRMGKSNEEDYDICESNIATPEE 1176
Db 1513 PINKNEHVSSITLNSKQDILTONGCNFVLYLLSFL-----SNKTSRKCVCANALY-- 1565
Qy 1177 NNYLNKLS-RIFNEOVET-----NISDYSEYNYNEKMY 1210
Db 1566 NSVLLQMPILRFRHNNLKTGYGNYGIRYDGIYKIANAFVNDYSTSEYKRDILY 1620

RESULT 27
US-10-087-013-9
; Sequence 9, Application US/10087013
; Publication No. US20040062769A1
; GENERAL INFORMATION:
; APPLICANT: Arthur Scherf
; APPLICANT: Louis H. Miller
; APPLICANT: Benoit Gamain
; APPLICANT: Dror I. Baruch
; APPLICANT: Pierre Buffet
; APPLICANT: Christine Scheidig
; APPLICANT: Jurg Gysin
; APPLICANT: Bruno Pouvellet
; APPLICANT: No. US20040062769A1utaka Fujii
; APPLICANT: Joseph Smith
; TITLE OF INVENTION: IDENTIFICATION OF THE DOMAIN OF
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ERYTHROCYTE MEMBRANE PROTEIN 1
; FILE OF INVENTION: (PFEPI) THAT MEDIATES ADHESION TO CHONDROITIN SULFATE A
; FILE REFERENCE: NIH176.001C1
; CURRENT APPLICATION NUMBER: US/10/087,013
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: PCT/US00/24195
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/152,023
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-087-013-9

Query Match 5.1%; Score 333.5; DB 4; Length 351;
Best Local Similarity 27.9%; Pred. No. 1.1e-10;
Matches 92; Conservative 46; Mismatches 131; Indels 61; Gaps 11;
Qy 442 IASQINVDLRGFCGNYKSNKESWNCGTFTNKFGCTCBPRQTLCLG-RTVLLHRGH 500
Db 9 ISQIEKQNH-----NCKTTEDAKWKCENTLGEDEGVCMPPRQNLGVHLYTKLNDSDK 64
Qy 501 EEDYKHELLGASIEAQLLYKYEK---DENAL-----CSIQNSYADLADIKG 548
Db 65 EEDLRAPAKSAAEFTLLRQYNSKNVEDDKILHRDMPPEPFRSMFYTFGVDRLCLD 124
Qy 549 SDI---IKDY-----YGKKEENLNKVKDKRNEESLKIIFREKWDENKENVK- 595
Db 125 TDISEKADHDVTAKKKTAVFQKIGSKTNGKKVLE--REGWKEYGSLSIWKMGLCAL 182
Qy 596 -----VMSAVLNKQKTCQDYDFQKIPQFLRWPFKEGDDFCEKREKIY 639
Db 183 SYNTETKMDGEVRYLMKLYKNDIKEVLEEFASRPFLRWTEWGEDFVNRKKELV 242
Qy 640 SFPSFKVEK-----KKKDCDEN-----TCKNKCSYKKWIDLKKSSEYKQVDKYTKDN 688
Db 243 SLKKKCDSCTLRNGTSNKTCDNENGCAKCTQCEKYKKWNERWKHYSSQKKKFLYKN 302

Qy 689 KMYNDIDEVONKEANVYLKESKECKDVN 718
Db 303 SATYNN--GLAVKEANSETYKNDPEVTEAN 330
RESULT 28
US-10-732-923-3342
; Sequence 3342, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 3342
; LENGTH: 2910
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-732-923-3342
Query Match 5.1%; Score 331.5; DB 5; Length 2910;
Best Local Similarity 18.5%; Pred. No. 1.9e-09;
Matches 289; Conservative 249; Mismatches 522; Indels 499; Gaps 70;
Qy 2 KGYFNIFLIPLIFLVN-IRINESIIGRTLNYR-----QDESSD-----ISRVNSPELN 50
Db 425 KSISSNVNFKTLFLFVNVLLKVSVHISG--LYTFMDILNYSDDLNNHINSNNNNN 482
Qy 51 NNHKTNIYDSY-----EDVN--NKLINS-----FVE 75
Db 483 NNNNNNNNDYFISVELRKLIESINIYQAVNAALYLFHYILHMLEITSQKNNMLIB 542
Qy 76 NKSVKKRSLSFINNKTYSYD--IIPSYSYRNDKFNSENEDSGNTSNPNFANTSEI 133
Db 543 NE-ISTIRNFIKIINKWKVDSNILINIVQYKN-----YPNTQNDNTY----- 585
Qy 134 SIGKDNKQVTFIQKTHLFCAGIKRKSIKWICRENSEKITVCVDPDKIQLCVAFNLSRL 193
Db 586 ---KSNKQLNIQOVENKKYKYNLINKE---LVHRNFE-----IRKILICKYN--NKII 630
Qy 194 ETMEKPKEIFLISV-----NTEAK---LLYNKVEGKDPs----- 224
Db 631 KWMKKIHLNLYIMHVYKSYKNDQNGYINBEEKNETKTETKNEKQNETKREINNEKQNETKR 690
Qy 225 IFCNELRNSFSDFRSSFIGDDMDPGGNTDR----- 254
Db 691 EINNETKTERNDERDERNNERNERNERNERNERNERNERNERNERNERNERNERNERN 750
Qy 255 -----VKGYINTKFSDYKKEKVEKLN-----NIKKEWKEKQKANLWNHMIVNH 298
Db 751 YQTPYCCNITYTKDKIENKKNYNIPEVESLSFFHCCNINITYYKKEIDINNMKLIY 810
Qy 299 KGNISKECAIIPAEPEQIINLWIKEMNENFLMEKRLFLNIKD-----KCYE-----NKKY 348
Db 811 LSNLL-----TKYQNSEFIKROMSYFKBINLFFDKIDYDILLCYNNMRNLNKED 863
Qy 349 EACFGCRLPCSSYTSFMKKSXTOMEVLTN-----LYKKKNSGVDKNNFLNDLFKKNNKN 403
Db 864 NINIPFCOL--NNSTCOLANNSTCELANNSTQNLNFTYNNKDLCSKEIDELLSDSMSKYNIN 921
Qy 404 DLDDFFKNEKEYD-----DLCDCRYTATIIKSFLNGPAKNDVDVIA----- 444
Db 922 -VDKIMKHYNTYNDILLKSNMKNVDMLLLLSNNIYNKIMSTTYNNFDIIFKIMKI 980
Qy 445 -----QINVNDLRGFCGNYKSNKESWNCGTFTNKFGCTCBPRQTLCLGRTYLLHRG 499
Db 981 FVHVIKIFVKDYTML---YINNQGKINESNEDHN-----LNKNKDSHID 1022
Qy 500 HBEDYKEHLLGASIEAQLLYK- -YKEDENALCSIQNSYADL--ADIIGSDIHKDY 555

Db 1023 NNIEGENFLANQKLIIDEVEKRVNIIKVELENEKNIINNIKEBQNEKIKYNDIKTNL 1082
QY 556 -YKQMEENLN-KVNDKKNESLAKIFREK---WWDENKE-----591
Db 1083 EREKKINMNIIELEKERKINDENMIKPEKKNQMYDINDKELENEKQNKIYIWEIQNIK 1142
QY 592 NVKQMSAVLNK-----ETCDYDFQKIQPOFLRMFKWGGDFCEKREKIKYSPFSKV 646
Db 1143 NINSDMNFALDEKSVQDQKCIDFEKLNK------YEDTCKELNNKQVLYEQMNT 1192
QY 647 ECKKDCDENTCKNKSSEYKWDLKS-----EYEQVDKVTKD-----686
Db 1193 QLYKEKINDEI---CTELEKIKNIINNIQINLENEKKTNEQKQIEKQKMKIKEMVIQ 1249
QY 687 -KNKQMYDN--IDEVNKEAN---VYLKEKCKQVNDKIFNPSNPEYEDMCKKCD 740
Db 1250 LEKEKINNEIITOLENEKLNKIY-----EDIEKKNKLSLEENYEQKQKINE 1301
QY 741 I-----KYLNE---IKYPTKTHDIYDITFDGTPGISINANINEQSGKDTST 790
Db 1302 MIQLEKEKIINNEIITOLENEKSVKINTLENIK-----KMDENMEKLNKEQKINN 1356
QY 791 G-----NSTSPSVHESPEAAINVE-KLSDGESSSETRI-----827
Db 1357 DLQDLENEKNEKDHINKEFEKIKNIQIKQNDDEKKTNEISLVNVEIDRLYKINEHVLL 1416
QY 828 -----LDNDPSVTVNNEVHDASNTQGSVNTSDI-----858
Db 1417 NLEBEKKIDQFTIDLENERKRNHIQLHPDDEKKNQNDQYEDLLKERKKNKLQILIYDD 1476
QY 859 -TNGHSESSL---NRTTNAQDIKI---GRSGNEQ-SDNOENSHSDNSGSLITIGQVPS 909
Db 1477 GKNNISQNLNLENERTRNDKILLDQEKKQNEQINDLENERKNNQNLQNLNEEQK 1536
QY 910 EDNTQNTYDSQ-NPHRDTNALASPSDDKINIEGFDSSRDSSENGKGTSTNTHDVRT 968
Db 1537 KEQLVSYEBQKINHOLENELOKQRTYK-KMAKPERKFLMN-----TNDTQIKOT 1590
QY 969 N-----IVSERRVNSHDFI---RGMANNNAHQYITQI-----ENNGIIRQGESAG 1013
Db 1591 QQIITQKIIDTQKIIDTQKIDTSNNVQNDDEKHKVQDQDAESDNTFLEQLEK- 1649
QY 1014 NSVNY-----KDNPKRSNFSENHKKNI-----QEYNSRDTKRVREBI-----1052
Db 1650 KQVNDMIQIKDKKIDELNLEKEKEVNDKIIIQEYQK-IEHINELEKEKEI 1708
QY 1053 -----IKLSQKKNKNNSEYCTYSDERNSPSPGCSREERKLLCC 1093
Db 1709 NNNLIRIEKNEKNEQNLQIDTEKKNQMSIE--LDEKQWQHRRITELDEKKNW 1766
QY 1094 QISDYCLKYNFYSIHYNCIKSEIKSPYKCFKSGQSSIPYFAAGGILVIVILLSSA 1153
Db 1767 ENN-----IBLEDEKKNK-----INILLEENK 1790
QY 1154 SRMGKN-----EYDICE-SNIBATPEENNYLNKLSRIFNQEYQETNISDYSVYNEK 1207
Db 1791 NKIIQINSCLQKEKKEINELLNIQLEKKN-----IQELNIKLEDEKKNLEQ 1837

RESULT 29

US-10-732-923-22709
; Sequence 22709, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149

; SEQ ID NO 22709
; LENGTH: 3124
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-732-923-22709

Query Match 5.1%; Score 329; DB 5; Length 3124;
Best Local Similarity 20.1%; Pred. No. 2.8e-09;
Matches 246; Conservative 20; Mismatches 434; Indels 344; Gaps 60;

QY 2 KGYFNIFLPIPLIFLVNVRINESIIGRTLYNNRQDESSDISRVNSPELNNHNTNYDSD 61
Db 1736 KGRMKTLYLLVDI--LDVVRKGSGL-----NYSSS 1764

QY 62 ---YEDVNNKLNLSFVENSVKKRSLSFNNKTKSYDIIIPPSYVRNDKFNLSNEDN 118
Db 1765 NLLLSQLSGEANSIYEREDI-KEGSDII--KESRDII-----KESRDIIKISTN 1815
QY 119 SGNVSNNTFANTSEISIGKNDKQYTFIQRTHLFACGIRKSI-----KWICHENSEKIT 173
Db 1816 ISKSSSRNISKSSRSI-SDIKEGQIIDKEDLIFKINRMKNKIDSRYSKRIDKESRDKI- 1873
QY 174 VCVPRKIQLCVANFLNSLETMEKEIFLISVNTAEKLLYNKEGKDPISFCNELRNS 233
Db 1874 -----SDKTNHVLDEVVHSDIHLN-----YEINNR-----CKMKG 1908
QY 234 FSDPRSSFIGDDMDFGNTDRVKGYINTKFSYDYKKNVEKKNLNKIKKWEKKNKANLNH 293
Db 1909 TNN-ENKLGIDIFNM--YDKIKYI--YKQYKSKSMENISFICH--YRNTKYKSDY 1959
QY 294 MIVNHKG-----NI-----SKECAIIPABE---PQINLW 319
Db 1960 LLLDNKESKPKRNTSYVLESPLHLIGDIVDNNIKRKKKKEIKTIVSDMFTSPVN-- 2017
QY 320 IKEMNFNLEKRLFLNFKOCVENKYEACPGGRLPCSSYTSFMKSKTOMEVLTNL 379
Db 2018 IKEYNTN--EOER-----KGEIVGNLSYDKTKIC--PPIKTEGR-----IKCNK 2060
QY 380 YKKNQSGVDKNFL--NDLFKK-----NNKNDLDDFFKNEKEYDLCDCRYTATIKS 430
Db 2061 IKKEKKEYNPNFLYNDYSSSPKYGDNENNFKYIRERKDFOKKFD--HPNPNFSK 2118
QY 431 FLN--GPAKNDVDIASQINVDLRGFCNYSKNEKSNWCTGTFNKFPGTCPPRRQTL 488
Db 2119 FLHNTYPMNK-NKNNKNNVNRREYPNYTSKDGVSYNFLSDLSFSSDN-----2170
QY 489 CLGRTYLLHRGHEDEYKHLGLASIIYEAQLLYKKEKDNALCSII-----QNSYA 540
Db 2171 -----YSSDNEYSDDSEK-----YKRVKKNKIIKFDLFTKIYIEKRLQMNND 2219
QY 541 DLADIIKGSDDIKYQKQWENLNKVNKDKKNEBSLAKIFREKWDENKENVKQMSAV 600
Db 2220 -----YKKG-----GKLNKNGMERNTKYKNVN--BITKMKYFVNNENRDHEV---2261
QY 601 LKNKE--TCKDYDKFQKIQPOFLRMFKWGGDFCEKREKIKYSPFSKVBC-----KKDCDE 655
Db 2262 --NKEDISKMOKY-----FLHISKHKEQIEDKKTTHYFKN--VECVYPAGNNINH 2312
QY 656 NTKCNKCSYKWIIDLKSEYEKQVDKYTKDNKKNMYNDIDEVKNKEANYLEKSKCK 715
Db 2313 NPSREKRYKS--INL-----YD-HLDEQSKIKGKKYFNKDKELIGSINKQTERKPKKN 2365
QY 716 DVNFDKIFNESPNEYEDMCKDEIKYLNELIKYPTKTHDIYDITFDGTPGTGISIN 775
Db 2366 KKNENK-----KDKKIRMTNNTKKEKHSNIIISVE-----EQN 2400
QY 776 ANINEQSGKTSNTGSETSDSPVSHESPEAAINVEKLSGDSSESSETRGILDINDPSV 835
Db 2401 MHNNSLKKEVNFQKNE-----EYLNANTNCSLIGKEWEEDVYEFHNNIYNNQTSY 2456
QY 836 TNNVNEVHDASNTQGSVNTSDITNGHSSSLNRTTNAQDIKIGRSGNEQSDNQENSHS 895
Db 2457 SDDIN---NTTKLKGNNNTNDISKNGKNLKG-----KISFFSMNNKYHESEIMNE 2506

QY 896 SDNSGLTIGQVPSDNTQNTYDSQNPDRTPNALASIPSDDK----- 938
Db 2507 EDKNMNLNLTQ-----SQIINKDKYNYFTCHPSLKKKKSQVFTKINNLFKNYFK 2554
QY 939 --INIEGFGDS-----SRDSENGRG-----DTSNTHTDVRTNIVSERRVNSHD 980
Db 2555 SIDVHEKFGSKFKPHSKDSDDIKGNKKISKRYNNNNNNNSYNSNIDSGK--YSHN 2612
QY 981 FINGMANNNAHQYITQIENNGIIRQOESSAGNSVNYKONPKRSPFSSSDHKKNIQY 1040
Db 2613 NKKNHHNNKKYHH-----NN-----NKVHHNNKKYHH--QNNYKKGHHNS 2654
QY 1041 NSRDTKVRREILKSKONCKNNEYSMEYCTYSDERNSSPGSPCRERKLC--CQISDY 1098
Db 2655 NSR-----VMSLGSKEKXENVDAIYQFDNYD-----KKLLKKLTSLNLQNK 2698
QY 1099 CLKYFNFPYSIEYCNKISEKSPY 1123
Db 2699 NVKNFNF-----YKFNDELEEE-EY 2719

RESULT 30
US-10-087-013-11
; Sequence 11, Application US/10087013
; Publication No. US20040062769A1
; GENERAL INFORMATION:
; APPLICANT: Arthur Scherf
; APPLICANT: Louis H. Miller
; APPLICANT: Benoit Gamain
; APPLICANT: Dror I. Baruch
; APPLICANT: Pierre Buffet
; APPLICANT: Christine Scheidig
; APPLICANT: Jurg Gysin
; APPLICANT: Bruno Pouvelle
; APPLICANT: No. US20040062769A1utaka Fujii
; APPLICANT: Joseph Smith
; TITLE OF INVENTION: IDENTIFICATION OF THE DOMAIN OF
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ERYTHROCYTE MEMBRANE PROTEIN 1
; TITLE OF INVENTION: (PFEMP1) THAT MEDIATES ADHESION TO CHONDROITIN SULFATE A
; FILE REFERENCE: NIH176.001C1
; CURRENT APPLICATION NUMBER: US/10/087,013
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: PCT/US00/24195
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/152,023
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-087-013-11

Query Match 5.1%; Score 327.5; DB 4; Length 308;
Best Local Similarity 27.8%; Pred. No. 2e-10;
Matches 85; Conservative 46; Mismatches 118; Indels 57; Gaps 10;
QY 455 GCNYSKNSWNCWCTGFTWKFPCTCEPPRPTLCGLRTYLLHRGHE-----EDYKEHLG 510
Db 10 GCMNPK-ESYPDWCKGNIDNSHGACMPRRQRLCVRD---LTQGGIRKPEPILITKFN 65
QY 511 ASIYEAQLLYKYKEKDENALCSI-----IONSAYDLADIIGSDIHKDYGK 558
Db 66 CAAKETHFAWHKYKDNVNAENELKSGKIPGFRKQMYTFGDFRDIFFCTDLSRCYIK 125
QY 559 KMEENLVNKKKQNE-----ESLKI PREKWDENKENVKMSAVLN-----K 604
Db 126 DTSQTKSLKGDQATTEKGDTHDDNKKL--QEWTTIHGPKIWEGLMCAITNGLSESEK 183
QY 605 ETCKDY-----DKFOKIPQFLRWFKWGGDDFCRKRKEKIYSFSPKVECK 649

Db 184 NILQDYSYNKLNNAEKDCCLCFASKPQFLRWYVWSDEFCHCRKKLEKVEDVICAK 243
QY 650 -----KKDCDENTCKNKCSEYKKWIDLKXSEYEQVDKYTKDKKKKKYDNIDVQNKAN 704
Db 244 DYEGCKNNKSNNSCVKCKEYENYITGKKTQYESQSGKFWTEKQKK-PEYNSYSKDKAS 302
QY 705 VYLKEK 710
Db 303 EYLKOK 308

RESULT 31
US-10-732-923-22588
; Sequence 22588, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 22588
; LENGTH: 3127
; TYPE: PRT
; ORGANISM: Plasmodium falciparum 3D7
US-10-732-923-22588

Query Match 5.0%; Score 326.5; DB 5; Length 3127;
Best Local Similarity 19.9%; Pred. No. 3.9e-09;
Matches 246; Conservative 198; Mismatches 428; Indels 365; Gaps 61;
QY 2 KGVNIYELIPLIFLVNVRINESIIGRTLXNRQDESSDISRVNSPELNNHKTNIYDSD 61
Db 1736 KGMKTYLLVDI--LDDVVRKGESL-----NYSSS 1764
QY 62 ---YEDVNNKLINSFVENKSVKKRSLSPINNKTYSYDIIPPSYSYRNDKFNLSNEDN 118
Db 1765 NULLSQLGSEAVSYIYEREDI-KEGSDI--KESSRDI--KEDSRDIKEISTN 1815
QY 119 SGNNTNNTANTSEISIGKDNQYTFIQKRLTHLFCAGIKRKSJ-----KWICRESEKIT 173
Db 1816 ISKSSRNISKSSRSI-SDIKEGQIIDKEDLIFKINRMKNKIDRSYKRAIDKESRDKI- 1873
QY 174 VCVPRDKIQLCVANFLNSRLETWEPKEIPLISVNTAEKLLYNKNEKODPSIFCNELNS 233
Db 1874 -----SDKTNHVLDEVVRKSHDIHLN-----YEINNR-----CKMKGD 1908
QY 234 FSDFRSFIQDDMDFGNTDRVKGINTKFSYDYYKEKNVEKLNKIKKEWKEKNKANLWNH 293
Db 1909 TNN-ENKLGIDIFNM--YDKKIKYI--YKKNYKSKSMENISFIKH--YRNTKYKSDY 1959
QY 294 MIVNHKG-----NI-----SKCALIPABE-----PQINLW 319
Db 1960 LLLDNKSGSKFKRNTSYVLESPLHLIGDIVDNNIKRKKKKKEIKTIVSDDMFTSPVN-- 2017
QY 320 IKEMNENFMEKKRLPLNIK-DKCVENKKEACFGGCLRPCSSYSYTFMKSKTQMEVLTN 378
Db 2018 IKEYNTNEQERKKEIVGNLSYDK-----TKKI-----PFFIKTKEGRIKKKIEKKEK 2066
QY 379 LYKKNKSGVDKNNFL-NDLFKK-----NNKNDLDDFFKNEKEYDDLCDCRYTATIIK 429
Db 2067 KEKKENN---NNFLYNDYSSYSSPKYGDNNENFVIKIRERKDFQKKFD--HPNPNFS 2120
QY 430 SFNL--GPAKNDVDIASQINVNDLRCFGCNYSKNNKSWNCTGTFTNKPCTGCTEPPRQT 487
Db 2121 KFLHNTNPMKN-----KMKKK--NNKNVRNREYNPNYSSSKD-- 2156
QY 488 LCLGRYLLHRGHEEDYKEHLKASIEYEAQLLYKYKEKDENALCS--IIONSAYDLADI 545
Db 2157 ---GVSY-----NFLSDSLFSSD---NEYSSDNEYSSDSEKYYKKFKKKNKI 2198

```
QY 546 IKGSDIHKDY------GKMEENLKNVKKRNEESLKIPREKWDE 588
D 546 IKGSDIHKDY------GKMEENLKNVKKRNEESLKIPREKWDE 588
D 2199 IKPDDLFTKIYIKKRLLOMNYDVKGKGLKNKGNERNKTKYKNV--EITMKYFVN 2256
QY 589 NKENWKMVAVLKVK--TCHDYDKFOKIPQPLRFKMGDDFCCKRKEKIYSPESFKVE 647
D 2257 NENRDHEV-----NKEDISKSMOKY-----FLHISKHKEQIEDKCKTKHYFHN--VE 2303
QY 648 C-----KKKDCDENTCKNCKSVKWKIDLKKSEYKQVDKTKDKNKKMYDNIDEVKNKEA 703
D 2304 CVYPAGNNINHNFRNEKRYIS--INL-----YD-HLDEQEKIKGKKYFNKOKELIGSI 2356
QY 704 NVYLKESKCKOVNPDCKIFNESPEYEDMCKDEIKYLNBIKYPKTHDIDYIDITFS 763
D 2357 NKQTERPKKKKNKNIENK-----KDKKKIRMTNKTKEKHSNLSIISVE--- 2400
QY 764 DTFDGTGTPISANINQOQSOTNTGNSSTSPVSHPESDAAINVEKLSGDESSSE 823
D 2401 -----EQNMHNHNSLKKKEVNFCKNE-----EVLNFRANTNCSLGIKEMEEDVYEFH 2447
QY 824 TRGILDINDPSVTNNVEHVDASNTQGSVNTSDITNGHSESLNRTTNAQDIKIGSGN 883
D 2448 SNNIYNNQTSYSDIIN---NTTKLKGNGNNTDISXNKGKNGKLGK-----KISPFM 2497
QY 884 EQSDQENSHSSDNGSLATIGQVPSDNTQNTYDSQNPHRDTPNALASLPDDK----- 938
D 2498 NKKYHSEIEMNEEDKNKMLNTQ-----SQIINKDKNYFTHCPSLKXKSVF 2545
QY 939 -----INEIEGDS-----SRDSNGRG-----DTSNTHDVVRT 968
D 2546 TKINLNFKNVFKSDVHEKFGSKFKFHSKSDDIKGNKNNKISKRYNNNNNNNNYS 2605
QY 969 NIVSERVNSHDFTRNGMANNNAHOYITQENNGIIRGQESAGSNVYKDNKRNSFS 1028
D 2606 NIDSGK--YSHNNKGNHNNNNKYYHH-----NN-----NKYHHNNNNKYYHH-- 2645
QY 1029 SENDHKNIQYNSRDKRVREIILKSKQNKCNBYSMECTYSDERNSSPGPCSEER 1088
D 2646 QNNYKHHHNSNR-----VMSKGEKTEKNENVYAYQFDNYD-----KKLL 2689
QY 1089 KCLC--COISDYCLKYNFYSIEYNCISKISPEY 1123
D 2690 KCLTSLNLQNLKKNVKNFMF---YKFNDELEB-EY 2722

RESULT 32
US-10-153-273-10
; Sequence 10, Application US/10153273
; Publication No. US20020169305A1
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; Chitnis, Chetan
; Miller, Louis H.
; Peterson, David S.
; Su, Xin-zhaun
; Welles, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/153.273
; FILING DATE: 21-May-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/210.288
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Fuller, Michael
; REGISTRATION NUMBER: 36,516
; REFERENCE/DOCKET NUMBER: NIH121.1FWDV1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 700 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-153-273-10

Query Match 5.0%; Score 325; DB 4; Length 700;
Best Local Similarity 24.4%; Pred. No. 7.6e-10;
Matches 120; Conservative 70; Mismatches 161; Indels 140; Gaps 22;

QY 478 GTCEPPRRQTLG-----LGRTYLLHRGHEEDYKEHLGASIEAQLLKVKY-----KE 525
D 8 GACAPYERLHLDYNLESIDTTSTTHKLLLE-----VCMARKYEGSNTHYTHQHTNE 62
QY 526 KDNALCSIIQNSVADLADIIKGSDDIKY-----YKKMEENL-----NKNVKKRNE 575
D 63 DSASQLCTVLARSFADIGDIVRGKDLVLDYDNKEQKQKLEQKLDIFKKIHKDVMKTN 122
QY 576 ESLKIP-----REKWDENKENYWK-----VMSAVLKNKET----- 606
D 123 GAQERYIDDAKGGDFPOLREDNMTSNRETVMKALICHAPKEANYFIKTACNVGKTNGQC 182
QY 607 -CKDYD---KFOKIPQFLRFKMGDDFCCKRKEKIYSPESFKVECKK-----K 651
D 183 HCIGGDVPTFYDVPQYLRWFEEWAEDFCRKKKGL---ENLQKQCRDYEQLYCSNGY 239
QY 652 DCDE-----NTCKNCKSEYKKWIDLKSEYKQVDKTYKD-----KNK 689
D 240 DCTKTIYKGLVIGEHCTNCSVMCRMYETWIDNQKKEFLKQKRYETISGGSGKSPK 299
QY 690 KMYDNIDEVKNKEANYLYKESKECKDVNFD-----KIFNESPEVEDMC-----K 736
D 300 RTKRAARSSSSDDNGYESFKYKGLKEVGQDVDFLKLNLK-----EGICQKQPVQVNE 354
QY 737 KCDEIKYLNBIKYPKT--KHDIYDIDTFDGTGTPISANINQOQSOTNTGNSSE 794
D 355 KADNVDTNE-KYKTFSTRTEICEPCPCWGLEKGGPPKW-----KGDKTCSAKTK 405
QY 795 T-----SDSPVSHPESDAAINVEKLSGDESSSETRGIL-----DINDPSVTNNVNE 841
D 406 TYDPKNITDIPVLPYDPSQQNLIKVKYKNCFKGAPGGGQIKKWKQCYDYDEHRPSSKNNNC 465
QY 842 VHDASN--TQ 850
D 466 VEGTWDFKFTQ 476

RESULT 33
US-10-732-923-3353
; Sequence 3353, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
```



```

; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-087-013-10

Query Match
  4.9%; Score 319.5; DB 4; Length 311;
Best Local Similarity 30.0%; Pred. No. 5.7e-10;
Matches 94; Conservative 42; Mismatches 112; Indels 65; Gaps 15;

Qy 455 GCN--YKSNNEK--SWNCTGTFNKFPGCTCPTCPPTTCLGRTVLLHGHBEEDY--KEHLL 509
Db 7 GCNKYAGDKYPCWDCNSQIHITHNGACMPPRQKLCVSLGTWDRKAIEIRTEFI 66

Qy 510 GASIEAQLLYKVKEDNALCSI-----IONSADLADIIGSDIHKDY- 556
Db 67 KSAAIETHFADRYKEDNGEAEALKNIGNIEGFKRQMYTFGVDYRDIFFGRDISTHAYI 126

Qy 557 ---GKMEENLNKYNKDK---KRNEESLKIFREKWDENKENVKWSAVL-----K 602
Db 127 SGVSPKVITILEKENDAKYAAKQNSNELL--DDWDDQHGKDIWEGMLCALTHKISDEEK 184

Qy 603 NKETCKDY-----DKFOKIPOLRFWFKWGDGDFCEKREKIYSPSPKVE 647
Db 185 KKEIKNKYSYKLNESPKGSKNVEDFAKQPLRFWFGDFEFCQREK---BAKVYS 241

Qy 648 CK-KKCD--EWT-----CKNCKSEYKWIIDLKSEYKQVDKYTKDK--NKGMYNDIDE 697
Db 242 CSDAKYDGCNTKSNASCVSACKVYEDYTKKVEYTKQKGFDAEKITDKGEYGF-- 299

Qy 698 VKNEANVYLKEX 710
Db 300 -STKDASEYLLKK 311

RESULT 35
US-10-732-923-3340
; Sequence 3340, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 3340
; LENGTH: 1939
; TYPE: PRT
; ORGANISM: Plasmodium chabaudi
US-10-732-923-3340

Query Match
  4.9%; Score 318.5; DB 5; Length 1939;
Best Local Similarity 19.9%; Pred. No. 6.1e-09;
Matches 282; Conservative 216; Mismatches 523; Indels 393; Gaps 68;

Qy 18 NVIRINESIIG----RTLYNRQDESSDISRVNSPELNNHKTNIYDS---DYED----- 64
Db 208 NLEKINKVIEKENNLRELKPEKMEKNEI--IESLDGTINDKKNAYEKLKLEISFBEKRWI 265

Qy 65 --VANKLI---NSFVKNKSVKKGSLSPFNNTKSYDIIPPSYVRNDKFNLSSENEDNS 119
Db 266 EMLDSKLIKEENFA-NKQAKLEKENETIIIEKLKIDIESREKDFPKSKEKPSM--ENELNT 323

Qy 120 GNTN-SNNFANTSISIGKONQVTFQKTHLP-----ACGKRSKI- 161
Db 324 LKSDLSKNACQMEYVYKLEIKDLSQSLVEKEREIFEIKNEYDDKINNMMKELSLSSNDKGD 383

```

```

Qy 162 KWCIRENSEKITVCPDRKIQL-----CVANFLMSRLTMEKFEIFLISVNT 209
Db 384 NTVLHSEEEKINKLLKEKETELNEIHKYKNLEIETIKNELNEKEEELKNAKHTVEYN 443

Qy 210 ---EAKLLYNK---NEGKOPSPFCNBLRNSFSFRSFIQDDDFGNTDRVKG---YI 259
Db 444 LTKELIKLEKKTEDAKEGHKNEL--NELNQLSKLNK-----EKDNIKNENTPEL 490

Qy 260 NTKFSDYVYKKNV-----EKLNNIKKEM-WEKNKANLWHEMIVNHNKNI 302
Db 491 NDKLSSLNSEVNILNKQKOTLGNDIKTLNDLNNKNEINTSDNKNMKOKEDLAWLNEEM 550

Qy 303 SKECAIIPABEPQINLMIKEMNENFLMEKKRLFLNIKDKCVENKKYKACFGCRLPCSSY 362
Db 551 EGKCVVIDEIK-----KYKNEIFMLEEK-----LKEK-----ENY 581

Qy 363 TSFMKSKTQMEVLTN-LYKQNSGVDKNNPLN---DLFKQN-----NQNDLDDF--- 408
Db 582 ADL-----NDEISILRNSIYVKEKEFIEMKPEYENKINLFNKNPBEKKNIYENELNSLRK 637

Qy 409 PKNE---KEVDDI-----CDCRYTATI---IKSFLNGPAKNDVDIASQINVDLRGFG 455
Db 638 YDNEQGLIKQIDELNIQKLTKEEYKLOLYNDNMHMFASICTKIDMPYSENIGSDLVDFV 697

Qy 456 CNY-KSNNEKSWNCTGTFTNK-----FPQTCPPPRQTLCL 490
Db 698 TAYIKRDESSDANPDTTHKEMVAELEKRAHAIVAELEKHEKEIAKLGSHKEVVLRL 757

Qy 491 G-----RTYLLHRGHE-----EDYKHLGASIEAQL---LKYKYKE-----KQNA 530
Db 758 GEQHKBTIILEKHQDVVTKLGEQHKENIILKEEHKDVVTKLGDQVKEIAKLKEBHA 817

Qy 531 LC-----SIQNSYADLADIIGSDIHKDYKQWER-NLNKNVNDKKN 574
Db 818 VVVAELEKHKLGEGHKEMVDELEKRAHDFVEGLEKHEKTALEKHEKSEMMNEVEKRRH 877

Qy 575 EESLKIFREKWWDENK---ENVVMKMSAV-LKNKETCKDYDKFQK--IPQFLRFKXWGD 628
Db 878 ADFVEGLEEKHKAETAKLGEHREVVALEKHEKHEVALEKHEKHEIAKLEEGHKEVMA 937

Qy 629 DFCERKEKIYSPESFKVECKKDCDNTCKNCKSEYKWKIDL-----KKSE 675
Db 938 ELGEKHEVWAGLEA-----KHNLEEGHKEMVALEKHEKRAHDLVAVLEEQHAKBIIKGE 991

Qy 676 YEKQV-----DKYTKDNKKNMYNDIDEV-----KNKEANVYLKESKECKDVNPDKIFN 725
Db 992 EHKEVVAGIBEKYKVEAKLAEBHKVVTKLGEQHKHEIAKLEGDHKE-----VYN 1042

Qy 726 ESPNEYEDMCKKCEIKYLNKIKYPTKTHDIYDITFSDTFTGCTPISINANINEQOSGK 785
Db 1043 EVEKKNASLLNMLEE-NHKNEIMIKLEEHKESASDLVEKLY-----QKDE 1086

Qy 786 DTGNTGN-----SETSDS-----PVSHEPESDAAINVKLSGDESSESTRG 826
Db 1087 EVKSNKKIEBLTNVIKDLNDSIMCYKKQILEEVEKREYNEERINKLQVNE----- 1139

Qy 827 ILDINDPSVTNNVNEVHDASN--TQGSVNTSDITNGHSESSLNRTTNAQDIKLGSGNE 884
Db 1140 MKQWMDKILKEKEKEIKKLNKLSNYKVFETKENTYKNSVMVNE--NKEKIIY----- 1191

Qy 885 QSDNQENSSHSS--DNSG---SLTIGQVPSDNTQNTYDSQNPHRDTPNALASLPDSDKIN 940
Db 1192 DSVCKENISSEVGEKGNLKWTLSLKCKERNIPFSINDKNSESELVDTIKSA-----YIN 1247

Qy 941 EIEGFDSSRSSENGRGGTTSTNTHDVRRTNIVSERRVNSHDFIRNGMANNAHYYITQIE 1000
Db 1248 KIEMYKKEIB-DNGK-----NIEDLK-----NKILDLSNELINLENKMKVLTDEN 1291

Qy 1001 NNGHIROQESAGNSVNYKDNPKKSNFSESNH-----KKNIQEYNSRDTKREVEILKLS 1056
Db 1292 NN--LKKEIBIKONKLEKEKENTEILNLDNDIILKGEISEWKDEEKLTKENIKKN 1349

Qy 1057 KONKCNNEYSMEYCTYSDERNSSPGCSREERKKLCQISDYCLKYFNFYFYSIYVNCIKS 1116

```

Db 1350 DIEQINKEYI-----KEENLM-----KEN-ENINEVTSKLN 1381
Qy 1117 EIKSPYKCPKSEQSQIPYFAGGILVIVLLSSASRMGKSNEEYDIGESNI-EATFE 1175
Db 1382 QIEIEKMKLEELNKVE-----LLLAEREKRETNMSISNDKNIVENNILEDTDS 1429
Qy 1176 ENNYLNKLSIFRQEOETWISDYSEYNN-EKN 1208
Db 1430 KQNNLNK-----NVEDKTGDDINCEKN 1451

RESULT 36
US-10-732-923-8666
; Sequence 8666, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 8666
; LENGTH: 2110
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-732-923-8666

Query Match 4.9%; Score 318; DB 5; Length 2110;
Best Local Similarity 19.2%; Pred. No. 7.3e-09;
Matches 272; Conservative 185; Mismatches 468; Indels 492; Gaps 63;

Qy 33 NRODESDISRVSPENNNHKTNIYDSYEDVNNKLINSFVENKSVKKRSLSFINKT 92
Db 338 NNDEKSGDSIGEDDDNNHKSDDNNIDEDNN-----DHKSEDNIGEDDDNDE 390
Qy 93 KSYDIIIPSYRNDKFNLSNEDNSGNTNSNFANTSBSISCKDNKQYTFIQKTHLF 152
Db 391 KGGD-----NNIDENDN-NSDHKSEDNIDENDNSHQ-----DQEQFHETKDI-- 437
Qy 153 ACGIKRKSIIKWI CRENSEKTVCPDRKIQLCVANFLNSLETWKEKPEIFLISVNTAK 212
Db 438 ---IKNSSYEH-----DNK-----NYYN-----KTGEDYK----- 460
Qy 213 LLYNKGKDPISIFCNELNSFSDFRSFIDGDMDFGNTDRVKGYINTKPSDYKEKV 272
Db 461 ---SDKENYSPTRFNKLKKE-----KIDEYDTKLKI 489
Qy 273 EKLNNIKKEWKEKNKANLWNHMI VNHKGNISKECALIPABEPQINLWKEWENFLMEKK 332
Db 490 EKREENKN-YEKDEHEY-----ESDNYDKE-----KIN-----KKK 520
Qy 333 RLFL-----NIKDKCVENKYEACFGGRLPSSYTSFMKSKTQNEV---LTNLVKKK 383
Db 521 ELIILKNDIENDSDTSEHKRDS-----RSSQKQKCEKRRRIKDEYNLRRTKIASK 575
Qy 384 NSGVKKNFLNDLPKKNKNDLDDFFKNEKEYDDLDCCRYTATIKSFNGPAKNDVDIA 443
Db 576 PSS-DNNNSNDDNNNDNNNDND--NNDNDND-----NNDNDND----- 613
Qy 444 SQINVNDLRFGGNYKSNNEK-----SWNCTGTFTNKPFGTCEPPRQTLCLGR 492
Db 614 ---DNDNDNDNDNDNNHKNDDNDNDILTCNMDKHLTKIPPIIKATLDYQHAGL 670
Qy 493 TYLLHRGHEEDYKEHL-----LGASIEAQLLYKKEKDNALCSIIQNSYADLA 543
Db 671 HWLLYL-----YKNNINGILADEMGLKTLQCSLLSY-----LAYFINWGPHL 715
Qy 544 DIKGSDIK-----DYVGKMEENLNKV-----NKKD----- 570

Db 716 VIVPTSILINWELKRPFCPCFKILSYGNQNERYKRVGVFNKDSFHICISSYSTVVKD 775
Qy 571 -----KKR-----NEESLKIFREKWD-----ENKENVVKVMSA 599
Db 776 HLVFKEKRWKIILDEAHNIKNFNTKRWNIILSKRDNCLLITGTPLQNSLEELSLHF 835
Qy 600 VLKXKETCDYKFKQKIPOFLRFKWDGDFCEKREKIKYISFESFVKCECKKOD----- 654
Db 836 LMPNIPT-SHLD-----FKWFSDDL-----LAIEKSIHHSKELIDRLHTV 877
Qy 655 -----ENTCKNKCSEYKKWIDLKSEYKQV-DKYTKDN----- 688
Db 878 IRPIILRLKCKVEKEMPKYEHIIKCKLRRQOILYDEFINNNKQNTLNTGNYIGLMN 937
Qy 689 -----KMYDNIDDEVKN-----EANYVLKESKECKDVNFD 720
Db 938 ILIQLRKCVCNHCDFLTKYIQTYPYVLSIRYVPRFFILFEKYVA-----D 985
Qy 721 DKIFESPNEY-----EDMKCKDEIKYNEIKYPTKH---DIYDIDTSDTFGDTPI 772
Db 986 FYLILFLHNEFTSLGRDVTKETSPSKGFDLAHILTKHNTNELYDNHHSILYD----- 1040
Qy 773 SINANINEQSGKDT-----NTGNSSETSDSPVSHPESDAAINVEKLSGDESSETRGIL 828
Db 1041 --NNHISELYDNHHSILYDNHHSILYDNHSHK-----NYKHSNGYT 1083
Qy 829 DINDP--SVTNVNVVHDASNTQGSVNSTSDITNGH--SESSLNRTTNAODIKIGRSGNE 884
Db 1084 YPNDPINNMMNPSGFTKTSEQGQIVSHERDNNYHMDHNNNNLLSKEMVNSLRDDN 1143
Qy 885 QSDNQENSHSDNSGSLTIGQVPSDNTQNTYDSQNPDRDTPNALASLPDDK----- 938
Db 1144 SNNFYKSLTNNND-----QTSIHDNKQCDYNKL--CADTFNNINSIGNEEKSLNVL 1197
Qy 939 -----INEIGFSDSRDSENGRGTTCNTHTDVRTNIYSERRVNSHDFIRN 984
Db 1198 NEQNNNSKDNNDNNNDNNNDNNNDNNNDNNNDNNNDNNNDNNNDNNNDNNNDNNH 1253
Qy 985 GWANNAH-----HOYITQIENN-----GIIRGOESAGNSVNYKONPKRSNFSEN 1031
Db 1254 ---NNQHCVNDNDWPSDYPTNIINHRNAFLSLKLLNQ-----NPLN-NDNNNNNNNGN 1307
Qy 1032 DHKKNQIYNSRDTKRVREIILKSKONKNNEYSMEYCTYSDERNSSPOPCREERKKL 1091
Db 1308 NNINYNNRYNSRNSRSLNSPSSNTSKMS-FQLDF-LYTNSFINQDALCKN----- 1359
Qy 1092 CCQISDYCLKYNFYISIE-----YNYCIKSEIKSPEVKCFKSGQSSIPYFAAGGILVVI 1146
Db 1360 -----SFFVNIINIEDVHSYIYNSIYKPIPKILSFSDFELTEL-----NNNYDIL 1405
Qy 1147 VLLLSASRMGKSNEEY-----DIGESNI---EATFEENNYL----- 1180
Db 1406 SLVIDPYNY-KSYNEVLYKMEEGTLTNQOGLGDDNNKHIIYHKSTSENMTHMKRKTFI 1464
Qy 1181 ---NKLRSIFNQEVQETNI-----SDYSEYNYNEKNMY 1210
Db 1465 YKYNMFKVINNDTQYQNIFTDDTNNSYNSLEHNLW 1501

RESULT 37
US-10-732-923-3352
; Sequence 3352, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 3352

Qy	879	GRSGNQSDNQENSSHSD--NSGLFTIGQVPSBD--NTQNTYDSQ-----NPHRDTN 928
Db	1078	QNSNNK---FEELTNVTKOLINDSIMCYKKKISBDEVKREYNEBIEIQKLNKRLNEMKOKSD 1134
Qy	929	ALASLPSDD-----KINETEGFDSRDS--ENGRGDTTSTNTHDVRRTNIVSERRVNSHDF 981
Db	1135	KEIIKKEDEIKLKKLSINYKIFPEIKENTYNSKMLINENKEIIVDNICKE-KISESDI 1193
Qy	982	IRNG-----MANNAHHOYITOIEN---NGIIRGOBESAGNSVNYK 1019
Db	1194	EGKGKLNKMLLSLKCKERSVFSVDNKNENNELVDIVRNTYINKIEVYKEIKONNNKWE 1253
Qy	1020	D-NPKSNFSSENDHKHGIQEYNSRDTKRVREELIKLSKONKCNNEYSMEYCTYSDRNS 1078
Db	1254	DLRKKIIDLNSLANLKNVKNVLTTEENDNLURKEM--EMKQNKYNE-----NDE--- 1299
Qy	1079	SPGCSREERKKLCCQISDYCLKYFNPYSITYEYNCIKSEIKSPYKCFKSEGQSSIPYFA 1138
Db	1300	-----ILNLNDIILK-----LTKELSEWKKEKIKLEND----- 1327
Qy	1139	AGGILVVIVLLSSARMGKSNEEYDIGESNIEATPEENLYNLKLSIFNQ-EVQETNIS 1197
Db	1328	-----IEQINKEYKIKEENLMIKINES--INEVNNLKNQIEIEKKCKE 1368
Qy	1198	DYSEYVYNEK 1207
Db	1369	NLNTSVNNK 1378
RESULT 38		
US-10-304-095-6		
; Sequence 6, Application US/10304095		
; Publication No. US20030134275A1		
; GENERAL INFORMATION:		
; APPLICANT: Long, David M.		
; APPLICANT: Metz, Anneke M.		
; TITLE OF INVENTION: Telomerase Reverse Transcriptase (TERT) Genes		
; FILE REFERENCES: 47714-5009-US		
; CURRENT APPLICATION NUMBER: US/10/304,095		
; CURRENT FILING DATE: 2002-11-26		
; PRIOR APPLICATION NUMBER: US/09/417,485		
; PRIOR FILING DATE: 2002-06-14		
; NUMBER OF SEQ ID NOS: 49		
; SOFTWARE: PatentIn Ver. 2.1		
; SEQ ID NO 6		
; LENGTH: 2184		
; TYPE: PRT		
; ORGANISM: Plasmodium falciparum		
; FEATURE:		
; NAME/KEY: unsure		
; LOCATION: (330)..(335)		
; OTHER INFORMATION: Xaa at position 330 = Leu or Ile;		
; OTHER INFORMATION: Xaa at position 335 = Asp or Gly.		
US-10-304-095-6		
Query Match 4.8%; Score 308; DB 4; Length 2184;		
Best Local Similarity 18.2%; Pred. No. 2.7e-08;		
Matches 260; Conservative 220; Mismatches 432; Indels 518; Gaps 65;		
Qy	5	FNIFYLPIFLFLYNVIRINESIIGRTL-VNRODESSDISRVNSPELNNHKNHTVYSDYE 63
Db	716	FN-YFIMPLIRPFELTKBSQTLHKTIFFDRKIW-----NHFTKI----- 754
Qy	64	DVNNKLSNVFNKSVKKGSLSPFNKTKYSYDIIPPSVSYRNDKNSLSENEEDNSGNTN 123
Db	755	-SNFCLYHQIFRNKCLKGRNE-----PKMDYVQNMFNVKKGKGIKTN-- 796
Qy	124	SNNFANTSIEISGKNQKTYTFIOKTRHLFACGIKRKS-I-KWICRESEKITVCPVDRKIQ 182
Db	797	-----KYIFIKK-----MKKOSTNKCINNKFSK---CIPKKK--- 826
Qy	183	LCVANFLNSLETWEKFEIFLISVNTAEKLLYNQNEKQDPSIFCNELRNSFSDFRSSFI 242

Query Match	4.6%	Score 295;	DB 5;	Length 1647;
Best Local Similarity	18.9%	Pred. No. 1e-07;		
Matches	259;	Conservative 184;	Mismatches 471;	Indels 460; Gaps 60;

Qy	16	LYNVIRI--NESIIIGRLYNRQDESSDISRVNSPE--LNNNHKTIYDSDYEDVNNKLIN	71
Db	108	LYSCLYEENENVLKGYLHYKHNCNGEHWYKYEIYYNQLQDLPFGGIEJNEB---	164
Qy	72	SFVENKSVKCKRSLSFNNKTKSYDIIAPPYSYKNDKFNLSIENEDSGMNTSNPNFANTS	131
Db	165	--ELSKLKEMAKPYEIKNDLKCRNSFEMLINEBKINSILDYLANVEKKITNTN----	215
Qy	132	EISIGKONKQYTFIQKTHLPACGIKRSKIKWICRENSEKITVCVDPDKIQLCVANPLNS	191
Db	216	-LIIPKELK-YDLLQ-----YQKEGIYWMINQEMSNVKGIG-----	249
Qy	192	RLETMEKFEKIFILISVNTAEKLLYNKNEGKDPISFCNELRNSPDSFSSFTGDDMDFG-	250
Db	250	LADMGMGKTIQAITL-----ILCQKL-NKLKBIKQDERSDDHDKVGY	291
Qy	251	-NTD-----RVKGYINTKFSDYK-----ERNVEKLNNIKKEWERNKAN	290
Db	292	KNDDEDEKCKKCKKSHLHGDIKVEDILSFEERDTKHEKSIXEI-DIKKSY--KNSSSI	348
Qy	291	W--NHHVNVHNGNISKECAIIPAE-----EPQINLWIKEWENPLME	330
Db	349	LYCKNSTSDLLKVBESDCSVILIESDDTQNEKNHKNKVMAENKID-----TENLVN	402
Qy	331	KGRFLNLIKDVENKKYKACFGGCRLPCCSSYTSFMKSKTQMEVLNLYKKCKNSGVDKN	390
Db	403	KK--IKNIKKECVNKK-----IKNNNNNNKSNFKN	433
Qy	391	NFLNDL-----FKQNKNDLDDFPNKEKYEDDLCDCRYTATIKSFNLGP	435
Db	434	DFINKLKGOTLIIAPVAAVMWQKSEIEKFPDENILNVVYHG-----NSKIISD--EEL	485
Qy	436	AKNDVDIASQINVDNLRGFGCNTYKSNNEKSNWCTGTFTNKPFGTCPEPRQTLCLGTYL	495
Db	486	IKYDIVITTSVAVE-----VNFKK-----IVNKHKOPCE-----YC-GRLYL	521
Qy	496	LHRGHBEDYKEHLIG-ASIIYEAQLLKYYKEKD-----ENALCSIIQN	537

Db	522	--PNNLDIHKYFCGPTAVRTEKRLKCRKKKNKDTALVAMKKFDETVPVTPRNVLLLEIMAN	579
Qy	538	SYADLADIKSGSDI I KDYYGKMEENLKNVKDKKEBESLKI PREKWDENKENYKVM	597
Db	580	SKLEDEKKNTEIEN---KSLHNTTKYKGGKNNQIKQ---RNAQKDSN-----I	636
Qy	598	SAVLKNKBTCKDYDKFKIQQFLRWFKEWGGDDFCEKRKEKIYSPESFKVECKKDCDENT	657
Db	627	DVIVLSSDSCKE-----KSSSSSENS	646
Qy	658	CKMKCSE--YKMWIDLKSEYEKQVDKYT-----KDKNKKMYDNIDEVKNKEANVY	706
Db	647	VYSPLRKRTSSRIIDLNLGFDKETMEQVIESLSEGGKKNKNYKNVKNVSNKVNKS---	703
Qy	707	LKEKSECKDVNFDKIFNESPN-----EVEDMKCKDEIKYLEIKYKPT---KHDYD	758
Db	704	---VKNHRDVKNVVKMNLNPLVNDMDCKRHVLKIYVYNTTIOnteLEKLNMGSE	759
Qy	759	IDTFSOTFDG---GTPIS-INAMINEQSGKOTSNTGNETSDSPVS---HEPESDAAINV	812
Db	760	LKVVLITMGHIGFTKVVELINRLVSAKYIRKELYDENTETEBEKDTSGVHNEGDGDTSMGV	819
Qy	813	EKLGSBESSETGILDINDPVSVNNVNEVHDASNTQGSVNTSDITNGHSSSSLNR---	869
Db	820	TNQGKGDKFKERKIKKEDEKSSIEKKRKTVDVISVGGDVKR-----DMSLNKRRN	870
Qy	870	TTWAQDIKIGRSGNEQSDNQENSSHSDNSGSLITIGQVPESDNTQNTYDSQPHRDTPNA	929
Db	871	TTNR-----GTTWKIKGETRYASNGKGN-----NDODNNNSIYS-----	906
Qy	930	LASLPDDKINETEGFDSRDSNGRGDTTNSHTRRTNIVSERRVNSHDFIRNGMANN	989
Db	907	-----DAKSYTSSDESIOGDI CSKKKEPK-NYKSI RSQRKSYCKSKSLESE	954
Qy	990	NAHHQVITQIENNGIIRGOESAGSNVNYKO-----NPKESNFSSENDHKK	1035
Db	955	-----KKEESDNSYNTDNYDCDDCDDCDDYNNCDDNNYCDCCDSS	996
Qy	1036	NIOEYNSRDTKYVREBIIKLSKQKN-----NEYSMEYCTYSDBERNSS	1079
Db	997	N--SPDSYSDIIVRKS--RMSKEKKRENMKI FDESALHQIYWNRIILDEAHRIKNRNTS	1052
Qy	1080	PGFCSREERKGLCC-----QISDY--CLKYFNFYSEIYNCIKSEIK-----	1119
Db	1053	---TTQSILNLKCGYRWCLTGTPLQNRISLYSLRIFTEFPYAYFYCKKCKCLLIN	1109
Qy	1120	-----SPEYKC-----FKSEGOSSIPYFAAGGILVWTVLLL	1150
Db	1110	YEMRDNKYCYFNHRSRINHFNFKRLKPIQSPGYRGELSGMSYL--KNEVLDKILLRR	1168
Qy	1151	SSASRMKG-----SNEBYDIGSNTBATFEE--NYYNLKLSRIFN	1188
Db	1169	TKGERKSDINLPLIIKIRKDKLSKEKDFYESLYKOTSTQFTNYVNSNTVLHN	1222

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2005, 20:38:54 ; Search time 8 Seconds
(without alignments)
170.880 Million cell updates/sec

Title: US-10-677-980-2

Perfect score: 6481

Sequence: 1 MKGYFNIYFLIPLIFLYNVI.....VQETNISDYSEYNYNEKNMY 1210

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 8323 seqs, 1129788 residues

Total number of hits satisfying chosen parameters: 8323

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA_New.*

- 1: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	160.5	2.5	1315	7	US-11-077-550-141
2	160	2.5	2053	7	US-11-013-759-9
3	156.5	2.4	1169	7	US-11-077-550-20
4	151.5	2.3	1189	7	US-11-074-176-134
5	147	2.3	1420	7	US-11-077-550-110
6	147	2.3	2314	7	US-11-013-759-11
7	146	2.3	888	7	US-11-077-550-112
8	145	2.2	1992	7	US-11-013-759-3
9	145	2.2	1992	7	US-11-013-759-13
10	145	2.2	2047	7	US-11-013-759-4
11	145	2.2	2047	7	US-11-013-759-7
12	141	2.2	865	7	US-11-077-550-100
13	141	2.2	867	7	US-11-077-550-80
14	140.5	2.2	858	7	US-11-077-550-22
15	138.5	2.1	864	7	US-11-077-550-102
16	138.5	2.1	876	7	US-11-077-550-82
17	136.5	2.1	876	7	US-11-077-550-106
18	136.5	2.1	876	7	US-11-077-550-108
19	135.5	2.1	866	7	US-11-077-550-104
20	135.5	2.1	860	7	US-11-077-550-175
21	133.5	2.1	862	7	US-11-077-550-94
22	133.5	2.1	866	7	US-11-077-550-88
23	133.5	2.1	867	7	US-11-077-550-96
24	133.5	2.1	867	7	US-11-077-550-98
25	133.5	2.1	870	7	US-11-077-550-92

26 133.5 2.1 871 7 US-11-077-550-84
27 133.5 2.1 871 7 US-11-077-550-86
28 133.5 2.1 871 7 US-11-077-550-90
29 132.5 2.0 862 7 US-11-077-550-171
30 130.5 2.0 873 7 US-11-077-550-167
31 130 2.0 438 7 US-11-074-176-152
32 129.5 2.0 871 7 US-11-077-550-2
33 129.5 2.0 871 7 US-11-077-550-8
34 129.5 2.0 871 7 US-11-077-550-26
35 129.5 2.0 871 7 US-11-077-550-153
36 129.5 2.0 873 7 US-11-077-550-6
37 129.5 2.0 873 7 US-11-077-550-149
38 129.5 2.0 873 7 US-11-077-550-151
39 129.5 2.0 879 7 US-11-077-550-159
40 129.5 2.0 894 7 US-11-077-550-4
41 129.5 2.0 1013 7 US-11-077-550-18
42 129.5 2.0 1127 7 US-11-077-550-40
43 129.5 2.0 1129 7 US-11-077-550-42
44 129 2.0 866 7 US-11-077-550-32
45 129 2.0 876 7 US-11-077-550-66

ALIGNMENTS

RESULT 1

US-11-077-550-141
; Sequence 141, Application US/11077550
; Publication No. US2005024435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 141
; LENGTH: 1315
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-11-077-550-141

Query Match 2.5%; Score 160.5; DB 7; Length 1315;
Best Local Similarity 18.0%; Pred. No. 0.0041;
Matches 264; Conservative 178; Mismatches 457; Indels 565; Gaps 69;
Qy 32 YNRDESSDLSRVNSPELNHKTNIYDSVDVNNKLINSFVENKSVKKRSLSFINNK 91
Db 77 YLRDTSKDFLQTNVVKLFNRIKNNVAG---EALDKLIIN-----ALPYLGN- 120
Qy 92 TKSYDIIPPSYSYRNDKF-----NSLSEN---EDNSGNTNSN-----NFPAN 129
Db 121 --SYSLI-----DKFDTNSNSVSNFLRLQDPGSGATKTSAMLTNLIIFGPGVNLNKE 170

Qy	342	CVENKXYEACFGGCELPSSVTSFWKSKTKTOMEVLTNLYKKNSGVDKNPLN	-----	394
Db	952	-IHTTGTGA-----NTALQTFVTKKVDENDKADDTNATITVGKDGTS	991	
Qy	395	---DLFKKNNKNLDLDDFFKNEKEYDDLLDCRYTATIIKSLFNGPAKNDVDIASOINVD-	450	
Db	992	GKVNTLKLKGNGLD-----IKTDKOGTVTFGINTQSLKAGDS	1030	
Qy	451	--LRGFCNYSKNEKSNWCTGTFTNKF-----PGTCEPPRQTLCLGRVYLL	496	
Db	1031	TTLNNGSLSTKNTASNEBQIQVAGDGVKPFAMVNNGVVAGIDGTTRITRDEIGFTGT	1086	
Qy	497	HRGHEEDYKXHLGLCASIVYEQLLKYKKEKDENALCSIIIONSADLADIIGSDIIKDY	556	
Db	1087	-NGSLDKSKPHLSKOGI-NAGGKIITNIQSCE-----IAKNSH---DAVTGGKI---	1131	
Qy	557	GKKMBENLKNVKDKKKNESLSKIFREKWNDDENKENYKVMASVYLKNETCKDYDKPQKI	616	
Db	1132	DLKTELE-NKISSAKTAQNSLHFSVA--DEQGNF-----TVSNPYSSYDTSKTSDV	1182	
Qy	617	PQFLRWFKEWGDDCEKREKIKISFSEFKVECKKDCDENTCKNKCSB--YKKWIDLKKS	674	
Db	1183	ITP-----AGENGITTKVKNKVVVRVGIDQTKG	1209	
Qy	675	EYEQVDKYTKDNK-----XMYDNIDEVKNEKANYLKEKSECKEDVN	718	
Db	1210	--LTPKLTGVNNGKGIVINSQNGQNTITGLSNTLANVTNDKGSVRTTEOGNIKD--	1264	
Qy	719	FDDKI-----FNESPN-EYEDMCKKCDIEIKYLN-----BIKY--PKTKHDI	756	
Db	1265	-EBKTRAASIVDVLISAGFNLQGNGEAVDFVSTYDTVNFANGNTTTAKVTYDDTSKTSKV	1323	
Qy	757	YDIDTFSDTFEGDGPISi-----NANTNEQOSGKDTSENTGSETSDSPVSHPEPSDA	808	
Db	1324	YDYNV-----DDTTIEVKDKLGKVTITLTSTGTGANKKALNQATGDALYK---ASDI	1374	
Qy	809	AINVEKLSGD-----ESSETRGILIDINDPSV--TNNVNEVHDASNTQGSVSNTSDI-	858	
Db	1375	VAHLNTLSGDIQTAKGASQANNGAGYVDAGNKVIYDSTONKYQAQN-DGTVDKTKVEA	1433	
Qy	859	-----TNGHSESSLNRTTNAQDIKIGRSGNEQSDNOENS-----SHSSDN--SG	900	
Db	1434	KDKLVAQAQPTDGTILAQMNKVSINKEQVANDANKQGINEDNAFVKGLEKAAASDKTKNA	1493	
Qy	901	SLTIGQVPSBDNTQNTYDSQNPHRDTPNALASLP-----SDDKINEIEGTF	945	
Db	1494	AVTVGDLNAVAQPLTFAG-----DTGTTAKLGETLTIKGGQDITDKLTDNNIGVVAGT	1548	
Qy	946	DSSRDSENGEGDITSNTHDVRRTNIVS-----ERRVNSHDFTRNGMANN	990	
Db	1549	D-----GFTVVKLAKDITLNSNVNAGGTKIDEGSIFVD--ANGOAKANTPVLISANG	1597	
Qy	991	-----AHHQYITQIEN-NGILIRQGESAGNSVNYKDNPKRSNFSS	1029	
Db	1598	LDLGKRVISNVGKTKDTDAAVQQLNEVRNLLGL--GNDNADGNQVNIADIKDPNNGS	1655	
Qy	1030	END 1032		
Db	1656	SSN 1658		

RESULT 3
US-11-077-550-20
; Sequence 20, Application US/11077550
; Publication No. US2005024435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick

? APPLICANT: Wayne, Jonathan
 ? TITLE OF INVENTION: Recombinant Toxin Fragments
 ? FILE REFERENCE: 1581.0130004
 ? CURRENT APPLICATION NUMBER: US/11/077,550
 ? CURRENT FILING DATE: 2005-03-11
 ? PRIOR APPLICATION NUMBER: 10/241,596
 ? PRIOR FILING DATE: 2002-09-12
 ? PRIOR APPLICATION NUMBER: 09/255,829
 ? PRIOR FILING DATE: 1999-02-23
 ? PRIOR APPLICATION NUMBER: PCT/GB97/02273
 ? PRIOR FILING DATE: 1997-08-22
 ? PRIOR APPLICATION NUMBER: 08/782,893
 ? PRIOR FILING DATE: 1996-12-27
 ? PRIOR APPLICATION NUMBER: GB9625996.5
 ? PRIOR FILING DATE: 1996-12-13
 ? PRIOR APPLICATION NUMBER: GB9617671.4
 ? PRIOR FILING DATE: 1996-08-23
 ? NUMBER OF SEQ ID NOS: 179
 ? SOFTWARE: PatentIn version 3.1
 ? SEQ ID NO 20
 ? LENGTH: 1169
 ? TYPE: PRT
 ? ORGANISM: Clostridium botulinum
 ? US-11-077-550-20

	Query Match	2.4%	Score 156.5;	DB 7;	Length 1169;
	Best Local Similarity	18.0%;	Pred. No. 0.0062;		
	Matches 220;	Conservative 169;	Mismatches 415;	Indels 421;	Gaps 55;
Qy	252	TDRVKGINTKPSDYKKEKVEKLNLIK	---EWEKKNKALWNHMIWNHKGNI	SKCA	307
Db	40	TDRI-WIIPRYTFGYKPEDFNKSSGFNRDVC	EYDYLN	-----TNDKNI	87
Qy	308	IIPAEFPQINLTWIKENENFMLEKGRFLN	IKDKCVENKKYEACFGC	-----RLPCSS	361
Db	88	-----FLQTWIKLFNRIKSKPLGEKLEMI	INGIPYLGDRVPLEE	128	
Qy	362	YTSFMKSKSQMEVLNLYKKKSGVDKNFL	NDLFKQNNKNDLDD	DFPKKEYVDL	421
Db	129	FN-----TNI-----ASVTNKLISNPG	EVERKGGI	-----FAN	157

Qy	422	RYTATIISPLNGPAKND---VDIASQINVDLGRFG-----C-NYKS--NNEKSWNC	468
Db	158	-----LIIFGPGPVLNENETIDIGQNHFSREGFGGIMQMFPCPYVSFVNNVQB---	208
Qy	469	TGTFNKNFGPCTCPPRRQTLCLGRTYLLHGHBEDEYKEHLLGASIEYBAQLLKYYKPKDE	528
Db	209	-----NK-----GASIFNRRGYFSDPALILMHLEHILVHLGL-YGIKYDD-	246
Qy	529	NALCSIIQNS---YADLADIIGSDII-----KDYGYKWE-----ENL	564
Db	247	---LPVINEKKFPMOSTAIQABELYTFGGQDPSIITPSTDTSIYDKVLQNFPGIIVDLR	303
Qy	565	NKYNKDKCRNEESLKIIFREKWD-----ENKENVWKVMSAVLKNKETCKDYDKFQKIPQF	619
Db	304	NKVLVCISDPNININIVKPKDKYKVFVEDSEGYKSI-----DVESPDFLYKS	351
Qy	620	LRW-FKEW-GDDFCRKRKEIYSPSFVCEKKKOCDETCNKCKSEYKKWIDLKSEY	676
Db	352	LMFGFTETNIAENYKIKRASYPFSDSLPPVKIK-----NLLONEIYTIIEGFGNISDKDM	405
Qy	677	EKQVDKYTKDKNKQMDYNIDEVNKEANVYLKEKSE-----CKQVNFDDKIFNESPNE	730
Db	406	EKEYRGONKAINQOAY---EESISKEHLAVYKIQMCKSVKAPGICIDVDNEDLFPADKNS	462
Qy	731	YEDMCKKCOBIIKYLNEIKYPKTKHDLYDIDTFSDTFGDGTPISINANINEQOSGKQTSNT	790
Db	463	FSDLSKNERIETNGSVNIENDFPINELIIDLTL-----ISKIELPSENTE	510
Qy	791	GNSETSDSPSVSHBPESDAAINVEKLJGDESS-----SETRGILDINDPSPVNNVNEVHD	844
Db	511	LTDENVDPVYEBKQPA-----IKKIFIDENTIFQIYLSTOFP-LDIRIDILTSFSDALL	564

```
QY 845 ASNTQGSVNTSDITNGHSESSLNRTTNAQDIKIRSG-----NEQSDN-----QENSSHS 895
Db 565 FSNKVSFFSMDYI-----KTAN-----KVYEAGLFAGWVKQIVNDFVIEANKSNT 610
QY 896 SDNSG-----SLTIGOVPSEDTQNTYDSON-----PHRDTFNALASIPS- 935
Db 611 MDKIADISLIVPYIGLALVGNETAKGNFENAFIAGASILLBEIPBLLIPVVGAFLES 670
QY 936 --DDKINEIEGFSRDSERGRD-----TTSNT--HDVR----- 966
Db 671 YIDNKNKIITIDNALTKRNEKSDMYGLIVAOWLSTVNTQFVTKBGMKALNYQAAL 730
QY 967 -----RTNIVSERVNSHDFIRNGMAN--NNAHHQYITQIEN--NG----- 1003
Db 731 EEIKYRNIYSEKESKINIDFNDSKLNNEGINOQAINNPNFINGCSVSYLMKKMIMPL 790
QY 1004 ----IRGOEESAGNSVNYKDNPKRSNFFSNDHKNIQYNSR-----DTRKVRRETIKL 1055
Db 791 AVEKLLDPDNTLKKQNLNYIDENKLYLGSAEYKSKVNYKLYKTIMPFDLSIYNTDILI 850
QY 1056 SKQKNCNE-----YSMEYCTYS-----DERNSSPGPCSRERKX 1090
Db 851 EMENKYNSEILNLIILNRKYDNKLNLDLSGYGAKVEYVGVDELNDKNQPKLTSSANSKIR 910
QY 1091 LCQISDYCLYKYNFYIEYVNCYCKSEIKSPEYK-----CFKSEG----- 1130
Db 911 VTQONQIIFNSVLDVFSWF-----IRIPKYNDGIONYIHNEYTIINCMKNNSGWKI 964
QY 1131 -----QSSIPYPAAGGILWVILLSSA-----SRMGK 1158
Db 965 SIRGNRIIWLIDINGTKSVFFEYNTREDISEYINRWFFVTYNNLNNAKIYNGKLES 1024
QY 1159 SNEYDICE--SNIEATFENNLYN-----KLSRIFNOEVOETNISD-----YSEY- 1202
Db 1025 NTDIKDIREVIANGEIIFKLDGIDRTOFTWMKYFSIFNTELSQNSIEERYKIQSYEYL 1084
QY 1203 -----NYNEKNMY 1210
Db 1085 KDFWGNPLMYNKEYMFMNAGKNSY 1109

RESULT 4
US-11-074-176-134
; Sequence 134, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Klaenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McAuliffe, Olivia
; APPLICANT: Perill, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore
; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; PRIOR FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Lactobacillus acidophilus
US-11-074-176-134

Query Match 2.3%; Score 151.5; DB 7; Length 1189;
Best Local Similarity 19.2%; Pred. No. 0.012;
Matches 159; Conservative 138; Mismatches 290; Indels 243; Gaps 41;

QY 513 IYE--AOLLKYK-YKEKDENALCISIQNSVADLADIIKSGDIIDKDYKGMKEENLNKVNK 569
Db 160 IFBEAAGVLHFQKQKEAQNOL-----KKTQDNLIRIND 193
```

```
QY 570 DKCRNBESLKIIFREKWMDEKENVVWKGAVLKNKTKCDYDFKQKIPQFLRWFKWEGDD 629
Db 194 LVKELESRLPLENEQ-----SSLAKY-KFQK-----S 220
QY 630 FCEKRKEKYSFESFKVECKKKDCDENTCKNK-----CSEYKKWIDLKSEYEQV 680
Db 221 GUDKKLKLAFIEIENINOQREDIQKSAKNKILLAKLDDDEVKDSQAQAVTKRAEYKCLR 280
QY 681 DK--YTKDNKNKMYNDIDEVKNKEANVYLKESKECKDNVDFDKIFNESPNEDYEMCKKC 738
Db 281 DERDHTQNKLLKSLSEL--NASLQMAEQSRQ-----FDDATKEEYKNQYKQLKQNL 332
QY 739 DEIKY-LNEIKYPTK-----HDIYDIDTSDTDPGD--GTPISINANI-----NEQSGKD 786
Db 333 VOLKADLDELKKEKKLQDEQVVKIERGOLT--GELNEDPEELNKKLLDRIINNYMQLLOD 391
QY 787 TSNGTGNETSDSPVSHPEPESDAAINVEKLSGDESSSETRG-----ILDINDPSTVNNVNE 841
Db 392 QATNNQIIVNLNSDLRRSQADTTYQ---TGDVSKQLTDAQKOLEQRLEGKLLTDKRQK 447
QY 842 VHDA---SNTQGSVNTSDITNGHSESSLNRTTNAQ--DIKIGRSGNEQSDNOENS--- 892
Db 448 EQNAIVRINKQNNQNT-ELTN-----LRQVVNAERNELEKVEARHEALVNIQKRHEGY 500
QY 893 -----SHSDNSGSL-TIGQV-----PSEDNTQNT-----YDSQNPHRDTPNA 929
Db 501 YGVRNVNLHNDPAGVIGAVGSELITFPAELEAAMTALGGVQDLITTESRISARNAINK 560
QY 930 LAS-----LPSD-----DKINETEGDSSR-----DSENGRGDTTSNTH-- 963
Db 561 LKQNHGRATFPLDGLRQYGIPOSTVTTLKSVDGFRGIASDLVESKTDODITAAINYL 620
QY 964 -----DVARTNIVSRRVNSHDFI-----RNGMANNNAHVOYITQIENNGI 1004
Db 621 GSVVIVDTIDTAMSAVQVRNRYRIVTLGDGVISPGGSMGTGGQNRNSPLQTATEINQL 680
QY 1005 IRQOESAGNSVNYKDNPKRSNFSSE-----NDHKNIQYNSRDTKRVREIILKLSQNK 1060
Db 681 EK-QIKTLKQNLN-EDQDKLENLVDQSNKVNABEQDLQDALRETSQAINEAAISFQOQEK 738
QY 1061 -----CNNEYSMEYCTYSD-----ERNSSPGCSR--EERKKLCCOISDYCLK 1101
Db 739 EVKRLDANTLYKSRIKDRNDRIELKQIKKEANDKMLLTQKQEEQAKQNDLQDK-IK 797
QY 1102 YFNFYIEYVNCYCKSEIKSPEYKCFKSEGGSSIPYFAAGGILVIVILLSSASRMGKSNE 1161
Db 798 NFNLS---QRIQDELKSLDPK-----IAVVTNKLNLSSQENKKNH 836
QY 1162 EYDIGESNIR-----ATFENNLYNKLSRIFNOEVOETNISDYSEYVNE 1206
Db 837 QIDNSEKQIEDLTAKULTILAQDENSMMNQNTANLEKQKSTI-----EQKNNE 882

RESULT 5
US-11-077-550-110
; Sequence 110, Application US/11077550
; Publication No. US20050244435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; PRIOR FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
```


Qy	124	SNFPANTSEISIGKDNKQYTFIQRTHLFACGIIKRKSJK	-----HALSQGLANDTKTRAASIGDVLNAGFNLOQNGE	1220	
Db	1174	SNTLANTVNDGAG	-----		
Qy	171	KITVCVPDRKIQLCVANFLNSRL	---ETMEKFEIPLISV-NTEAKLLYNKEGKDSIF	226	
Db	1221	AVDFPVSYDTVDFIDGNATTAQVYDTSKTSKVYVDVNDKTIETVSDKKLGKVTTL	1280		
Qy	227	CNELRNSFSOPRSSIFIGDM	-----DFGNTDRVKGYINTKFSQVYKEKNVEKL	275	
Db	1281	TKTSANGNAIKESAADGALVKASDIATHLNTLAGDIQTAKGASQASSSASYVDAGNKV	1340		
Qy	276	--NNIKEMWEKKNKANLWNHMIWNHKGNI	SKCEAIIPAEBPQINLMIKEWENFLMEKKR	333	
Db	1341	IYDSTDKKYIQ	-----VNDKGQVDKNKEV	1364	
Qy	334	LFLNIDKCVENKKYKACFGCRLPCSSSYTSFMKSKTQMEVLTNLYKKNSGVDKNNFL	393		
Db	1365	---AKOKLVAQ	---AOTPDGTLAQMNVKSVINKEQVNDANKQ--GINEDN--	1407	
Qy	394	NDLFKNNKNDLDDF-FKNEKEYDDLCDCRYTATIIKSFLNGPAKNDVDIASQINVNDLR	452		
Db	1408	--AFIKLENAADKTKN	-----AAVTVGDLNVAQTPLTFAGD	1445	
Qy	453	GFGCNYKSNNEKSNWCTGTFTNKEPGCEPPROTCLGRTRYLLHRGHESDYKEHLIGAS	512		
Db	1446	-----TGTAKK	-----LGETLTIKGG	1462	
Qy	513	IYEAQLLYKYEKENALCSIIQNSYADLADIIKGSD	---IIKDYKGMKEENLKV	567	
Db	1463	-----QYDTNKLTD	-----NNIGVVAGTIDGFTVKLAKOL	1497	
Qy	568	NKDKRNEESLUKIPREKWDENKENVWKSVAVLKNKETCKDYDKFOKIPQFLRWFKEWG	627		
Db	1498	NAGGTRIDEXGISF	---VDANGQA--KANTPVL SAN	1529	
Qy	628	DDFCEKEKEKITYSPESFKVECKKDCDBENTCKNKSBYKKWIDLKXSEYKQVDKYTKD	687		
Db	1530	LDLGGKRISNIGA	-----AVDNDAVN---FKQFNEVAKT	1561	
Qy	688	NKKMYNDIDEVKNKEANVYLKEKSECKQVNFDDKIFNESPNEVEDCKKCEIKY	--L744		
Db	1562	-----VNNLNNQNSGASLPFVVTDANGKPI NGTD	-----GRQXAKIAGADGKYHANA	1610	
Qy	745	NEIKYPTKHDIYIDTFSDTFDGGTPI SIN	-----ANNEQOSGK DTS	788	
Db	1611	NGVPVDXGKPI TDADKLANLAHAGKPLDAGHQVVASLGGNSDAITLTNIKSTLPQIDTP	1670		
Qy	789	NTGNSETSDSPVSHEPESDAINVEKLSGDESS	--SETRGILDINDPSVTNNVNEVH--D844		
Db	1671	NTGNANAGO	-----AQSLPISLAAQOSNAASVKDVLNFGVFNLOTNH-NQVDFVK	1718	
Qy	845	ASNTQGSVNT-SDITNGHSESSL--NRRTNAQDIKIGRSGNEQSDNQENSSHSDN	---898		
Db	1719	AYDTVNFVNGTGADITSVRSADGTWSNITVNTALAATDDGCVLJIAKADGKFKYKADLMP	1778		
Qy	899	SGSLTIQGVSEDNQNTYDSONPHR	--DTPNALA-----SLPSDDK	938	
Db	1779	NGSLKAKGASDAKTPTGLSLVNPNAKGSGTGDAVALNNLISKAVFKSKDGTITTTTTSDDG	1838		
Qy	939	INEIEGFDS	-----SRDSENGRGDTTGN	---THDVRRTNIVSERRRVNSHDFTRNGWANN	989
Db	1839	IS-IQKDNSSITILSKDGLNVGGKVINVGKGTXTDAAV	---QQUNE---VRNLLGLG	1891	
Qy	990	NAHHQYITQIENNGIIRGQEEBAGNSVNYKDNPKRNSPSSEND	1032		
Db	1892	NA	-----GNMDAGCNQVNIADIKDPNPSGSSSN	1919	

```

; Publication No. US20050244435A1
;
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
;
; TITLE OF INVENTION: Recombinant Toxin Fragments
;
; FILE REFERENCE: 1581.0130004
;
; CURRENT APPLICATION NUMBER: US/11/077,550
;
; CURRENT FILING DATE: 2005-03-11
;
; PRIOR APPLICATION NUMBER: 10/241,596
;
; PRIOR FILING DATE: 2002-09-12
;
; PRIOR APPLICATION NUMBER: 09/255,829
;
; PRIOR FILING DATE: 1999-02-23
;
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
;
; PRIOR FILING DATE: 1997-08-22
;
; PRIOR APPLICATION NUMBER: 08/782,893
;
; PRIOR FILING DATE: 1996-12-27
;
; PRIOR APPLICATION NUMBER: GB9625996.5
;
; PRIOR FILING DATE: 1996-12-13
;
; PRIOR APPLICATION NUMBER: GB9617671.4
;
; PRIOR FILING DATE: 1996-08-23
;
; NUMBER OF SEQ ID NOS: 179
;
; SOFTWARE: PatentIn version 3.1
;
; SEQ ID NO 112
;
; LENGTH: 888
;
; TYPE: PRT
;
; ORG: Clostridium botulinum
;
; US-11/077-550-112

```

	Query Match	2.3%; Score 146; DB 7; Length 888;
	Best local Similarity	18.7%; Pred. No. 0.019;
	Matches	Conservative 138; Mismatches 344; Indels 308; Gaps 45
Qy	252	TDRVKGINTKPSDYKEKVKNLNITKK-----EWWEKKNKANLWNHVIWVHGHSKECA 307 : : : : : : : : : : :
Db	51	TDRI-WIPIERYTFGYKPEDFNKSSGFNRDVCYYPDVLN-----TNDKNI----- 98 : : : : : : : : : : :
Qy	308	IIPAEBOEINLIWKENNENFLMEKRLLFKDKCVENKKYEACFGGC-----RLPCSS 361 : : : : : : : : : : :
Db	99	-----FLOTMIKLFRNIRKSPGLGEKLEMLINGIPYLGDRAVPLEE 139 : : : : : : : : : : :
Qy	362	YTSFMKSKSQMEVLTNLYKKXNSGVDDNNFLNDFLKNNKNNDLDPPFKNEKEYDDLDCD 421 : : : : : : : : : : :
Db	140	FN-----TNI-----ASVTNKLISNPGBVERKKGI-----FAN----- 168 : : : : : : : : : : :
Qy	422	RYTATIISKPLNGPAKND-----VDIASQINVNDLRGCG-----C-NYKS--NNEKSWNC 468 : : : : : : : : : : :
Db	169	-----LIIFGPGEVLNENETIDIGIQHFASREGFGGIMQMCKPEYSVFNNVOE--- 219 : : : : : : : : : : :
Qy	469	TGTFTNKPGTCBPRRQTCLGRTYLLHRGHBEDYKEHLILGASIEYAQLLKYYKKEDE 528 : : : : : : : : : : :
Db	220	-----NK-----GASIFNRRGYSDDPALILMHELIHVHLHGJ-YGIKVDD- 257 : : : : : : : : : : :
Qy	529	NALCSIIQNS---YADLADIIGSDII-----KDYGGKME-----ENL 564 : : : : : : : : : : :
Db	258	--LPIVPNEKKPFMOSTDALQAEELYTFGGQDSPSIITPTDSIKSYDKVLGNFRGI VDR L 314 : : : : : : : : : : :
Qy	565	NKNVKDKRNESLIKIPREKWD-----ENKENVKNWSAVLKNKETCKDYDKFOKIPQF 619 : : : : : : : : : : :
Db	315	NKVLVCISDNINIINYKNPKOKYKEVEDSEGKYSI-----DVESFDPLYKS 362 : : : : : : : : : : :
Qy	620	LRW-FKEW--GDPCEKRRKEKISYFSBFKVECKKCDCTCNKCSYKKWIDLKKSEY 676 : : : : : : : : : : :
Db	363	LMFGFTETNAENVKI TRASYFSDSLPPVKIK-----NLDDNEITYIESGFNISDKDM 416 : : : : : : : : : : :
Qy	677	EKQVDKYTKDKNMKYDNIDEVNKNKEANVYLKEKSKE-----CKDVNFDDKIFNES 727 : : : : : : : : : : :
Db	417	EKEYRGONKAINKOAY---EEISEKHVLAVYIOCKSDDDDDKAPGICIDVNBELFFTD 473 : : : : : : : : : : :


```
RESULT 9
US-11-013-759-13
; Sequence 13, Application US/11013759
; Publication No. US20050249747A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Sasaki, Ken
; APPLICANT: Yang, Yan Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE
; FILE REFERENCE: 1038-921MIS:jb
; CURRENT APPLICATION NUMBER: US/11/013,759
; CURRENT FILING DATE: 2004-12-16
; PRIOR FILING DATE: 1999-07-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-11-013-759-13

Query Match      2.2%; Score 145; DB 7; Length 1992;
Best Local Similarity 17.6%; Pred. No. 0.054;
Matches 220; Conservative 161; Mismatches 430; Indels 438; Gaps 55;

QY 44 VNSPELNHNKTIYDSYEDVNNKLINSFVENKSVKKRSLSFINNKTSYDIIPPSYS 103
DB 527 VTAFTYNGVKTTELNSDGTGSKFSVKGSGTNNSLVTAHLASVLYNEVNRATDALSQFT 586
QY 104 YR---NDKFNLSNEDNSGNTNNTNFANTSEISI-----GKDNKQYTFIQRTHLFCAGI 156
DB 587 VKEEDDDANAITVAKOTTKNAGA-----VGLKLGKNG--LTVATKQGTVTIFGL 636
QY 157 KRKIKWICRE--NSEKIVCVPRKQIQOLCVANPLNSRLRLETMEKFEIFILSVNTEAKLL 214
DB 637 SQDGLTIGKSTLNDGLTVKDTNEQIQV--GANGI-----KFTNV----- 675
QY 215 YNKNEGKPSI--FCNELRNSFSRFSFIQDDMPGNTDRVKGVIYNTKFSYYKEKVE 273
DB 676 -----NGSNPGTGIANTR-----ITRDKIGFAGS-----DGAVDT----- 706
QY 274 KLANIKKEWKEKNKANLWNHMI VNHKGNISKECAIIPAEBPQINLWIKENWENFLMEKKR 333
DB 707 -----NKPYLDQDKLQGVNVIKTIWGINA----- 730
QY 334 LFLNIKDKCVENKYEACFGGCRULPCSSYT--SPMKSKTQMEVLTNLY-KKQNSGVVDKN 390
DB 731 -----GKAITGLSPTLPSTADQSSRNIELGNTIQDKDSNAASIN 771
QY 391 NFLNDLFKNNKNDLDDFFKNEKEVDLDCRYATATIKSPLNGPAKNDVDIASQINVD 450
DB 772 DILMTGNLKNNAPIDFVSTYDIDVDFANGNATTATVTHDTANTKSK-----VVYDVNVD 827
QY 451 LRGFPGCNYSNNEKSWNCTGFTTNKFPCTCEPPRRQTLCLGRT---YLLHRGHEE----- 502
DB 828 -----TTIHLTGDDNKKLGKVGKTKLTKNTSANGNTATNFVNSSDEDALVNA 874
QY 503 -DYKEHL--LGASIV-----EAQLKYKYEKENALCSIIQNSYADLADIKGSIIK 553
DB 875 KQIAENLNTLAKEIHTTKGTADTALQFTVKKVDE-----NNNADANAIT----- 920
QY 554 DYVGKMEENLKNYKDKKRNEESIKIFREK-----WWDENKENYKWSAVL-----K 602
DB 921 --VGQKUAN--NQNTLTLKGENGLNITKTDKNGVTFTGINTSLGKAGKATLNDGGLSIK 976
QY 603 NK-----ETCKDYDKFKQI-----POFLRW 622
DB 977 NPTSEQIQVADGVKFAKVNNGVVGAGIDGTTITRDEIGFTGTNGSLDKSKPHLSK- 1035

623 FKWGGDDFCERKEKIKYSPESFKVECKKDC-----DENT-CKKCKSEYKWKIDLKKS 674
1036 -----DGINAGGKITTNIQSGBIAQNSHDAVTTGGKIYDLKTELENKISISAKTQNSLH 1089
675 EY---EKQVDKVTDKNNKMYD-----NIDEVK----- 699
1090 EFSVADEQGNFTVSNPYSSYDTSKSDVITFAGENGITTKVNVGVVRVIGIDQTKGLTTP 1149
700 -----NKEANYLKEKSKCKEKNVDFDDKI-- 723
1150 KLTGVNNGKGVVIDSQNGQNTITGLSNTLANVTNDKGSVRTTEQGNIIKD--EDKTRA 1206
724 -----FNESPN-EYEDMCKKCKDEIKYLN-----EIKY---PKTKHDIYDIDTFS 763
1207 ASIVDVLSAGFNLQNGEAVDFVSTYDTVNFADGNATTAKVYTDYDTSKTSKVYDVNV-- 1264
764 DTFGDTGTPISI-----NANINEQSGKDTSGTNSSETSDSPVSHPEPSDAAINVEKL 815
1265 -----DOTTIEVKDKLGVKTTTLTSTGTGANKFALSNOATGDALVK---ASDIVAHLNTL 1317
816 SGD-----ESSSETRGILDINDPSV--TNVNVNEVDASNTQGSVSNSTSDI----- 858
1318 SGDIQTAKGASQANNSAGYVDADGNKVIYDSTDNKYVQAKN-DGTVDKTKEVAKDKLVAQ 1376
859 --TNGHSESSLNRTTNAQDIKIGRSGNEQSDNOENS-----SHSSDN---SGSLTIGQV 907
1377 AQTPTDGLTQAMNVKSVINKEQVNDANKQGINEDNAPVKGLEKAAKSNKTKNAAVTVGDL 1436
908 PSEDNTQNTVDSQNPHRDTPNALASLP-----SDDKINETEGPDS----- 947
1437 NAVAQPTLPFAG-----DGTGTTAKLGETLTIKGGQTDNKLTDNNIGVAGTDGFTVKL 1491
948 SRDSEN-----GRGDTTNTSHDVRTNIVSERRVNSHDFIRNGM-----AN 988
1492 AKOLTNLNSVAGGTKIDDKGVSVFVDSGQAKANTPVLISANGLDLGGKVISNVGKGTDT 1551
989 NNAHQVYIQTLEN-----NGIIRGOEESAGNSVNYKDNPKRSNPFSSEND 1032
1552 DAANVOQLNEVRNLLGLGNA---GNDNADGNQVNIADIKKDPNSGSSN 1597

RESULT 10
US-11-013-759-4
; Sequence 4, Application US/11013759
; Publication No. US20050249747A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Sasaki, Ken
; APPLICANT: Yang, Yan Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE
; FILE REFERENCE: 1038-921MIS:jb
; CURRENT APPLICATION NUMBER: US/11/013,759
; CURRENT FILING DATE: 2004-12-16
; PRIOR FILING DATE: 1999-07-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2047
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-11-013-759-4

Query Match      2.2%; Score 145; DB 7; Length 2047;
Best Local Similarity 17.6%; Pred. No. 0.056;
Matches 220; Conservative 161; Mismatches 430; Indels 438; Gaps 55;

QY 44 VNSPELNHNKTIYDSYEDVNNKLINSFVENKSVKKRSLSFINNKTSYDIIPPSYS 103
DB 527 VTAFTYNGVKTTELNSDGTGSKFSVKGSGTNNSLVTAHLASVLYNEVNRATDALSQFT 641
```


Db 737 LEBIIKYRNIYSEKESKNINIDFNDINSKLNENGINAIDNINNFINGCSVSYLMKMWIP 796
 Qy 1004 -----IIRGQESAGNSVNYKDNPKRNSPSSENDHKKNIQEYNSR-----DTKRVREIILK 1054
 Db 797 LAVKELDFDNTLKKNLNLYIDENKLYLIGSAEYKSKVNYLAKTIMPPDLSIYNTDIL 856
 Qy 1055 LSKQKCNKN 1063
 Db 857 IEMFNKYN 865

RESULT 13

US-11-077-550-80
 ; Sequence 80, Application US/11077550
 ; Publication No. US20050244435A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Shone, Clifford Charles
 ; APPLICANT: Quinn, Conrad Padraig
 ; APPLICANT: Foster, Keith Alan
 ; APPLICANT: Chaddock, John
 ; APPLICANT: Marks, Philip
 ; APPLICANT: Sutton, J. Mark
 ; APPLICANT: Stancombe, Patrick
 ; APPLICANT: Wayne, Jonathan
 ; TITLE OF INVENTION: Recombinant Toxin Fragments
 ; FILE REFERENCE: 1581.0130004
 ; CURRENT APPLICATION NUMBER: US/11/077,550
 ; CURRENT FILING DATE: 2005-03-11
 ; PRIOR FILING DATE: 10/241,596
 ; PRIOR FILING DATE: 2002-09-12
 ; PRIOR FILING DATE: 09/255,829
 ; PRIOR FILING DATE: 1999-02-23
 ; PRIOR FILING DATE: PCT/GB97/02273
 ; PRIOR FILING DATE: 1997-08-22
 ; PRIOR FILING DATE: 08/782,893
 ; PRIOR FILING DATE: 1996-12-27
 ; PRIOR FILING DATE: GB9625996.5
 ; PRIOR FILING DATE: 1996-12-13
 ; PRIOR FILING DATE: GB9617671.4
 ; PRIOR FILING DATE: 1996-08-23
 ; NUMBER OF SEQ ID NOS: 179
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 80
 ; LENGTH: 867
 ; TYPE: PRP
 ; ORGANISM: Clostridium botulinum
 US-11-077-550-80

Query Match 2.2%, Score 141, DB 7, Length 867;
 Best Local Similarity 18.6%, Pred. No. 0.036;
 Matches 181, Conservative 138, Mismatches 344, Indels 308, Gaps 45;
 Qy 252 TDRVGVINTKFSYDYKEKNVEKLNKK-----EWEKKNANLNMHNVHKGNIKSECA 307
 Db 46 TDRI-WIIPERYTGYKPFNKSGIFNRDVCYYPDYLN-----TNDKKNI----- 93
 Qy 308 IIPAEPPQINLIWIKENWENFLMEKKRLFLNIKDKCVENKYEACFGG-----RLPCSS 361
 Db 94 -----PLQTWIKLFLNRKSKPLGKLEMTINGIPYLGDRRPLBE 134
 Qy 362 YTSFMKSKTQMEVLTWLYKKNSGVKNPNFLNLFKNKNKNDLDDPFKNEKEYDDLDCDC 421
 Db 135 FN-----TWI-----ASVTNKLISNPGVEYKKG-----FAN----- 163
 Qy 422 RYTATIIFKFLNGPAKND-----VDIASQINVDLGRFG-----C-NYKS--NNEKSWNC 468
 Db 164 -----LIIFGPGVLNENETIDIGIQNHFPASREGFGIMQMKFCPEYVSFVNNVQE--- 214
 Qy 469 TGFTNKPFGTCBPRQTLGRTYLLHRGHEEDYKHLGASIEAQLLKYYKEKDE 528
 Db 215 -----NK-----GASIFNRRGYFSDPALILMHHLIHLGL-YGKIVDD- 252

Qy 529 NALCSIIIONS---YADLADIIKGSII-----KDYKQKOME-----ENL 564
 Db 253 ---LPIVPNEKKYFMQSTDAIQABELYTFGQDPSIITPSTDKSIYDKVQNFQFGIVDRL 309
 Qy 565 NKVNKKRNEESLKIIFREKWD-----ENKENVWKMVASLVKVKETCKDYDFKQIPQF 619
 Db 310 NKVLVCISDPNINININIKYKFKDYKFEVDESEKYSI-----DVESPDKLYKS 357
 Qy 620 LRW--FKEW--GDDPCEKREKIYSFESFKVECKKDCDENTCKNKCSEYKWKWIDLKSEY 676
 Db 358 LMGFTETNIAENYKIKTRASYFSDSLPPVKIK-----NLLDNEIYTBEGFNISDKDM 411
 Qy 677 EKQVDKYTKDKNKKMYDNIDDEVKNKEANVYLKEKSE-----CKDNFDDKIFNES 727
 Db 412 EKEYRGONKAINQAY---EESIKHLAVYKIQWCKSDDDDKAPGICIDVDNEDLFFIAD 468
 Qy 728 PNEYDMCKKDEIKYLYNEIKYPTKHDIYDIDFSTFGDGTTPISINAINNEQOSKDT 787
 Db 469 KNSFSDDLKSKNERIEYNTQSNYIENDFPINELILDIDL-----ISKIELPREN 516
 Qy 788 SNTGNSSETSDSPVSHPESDAAINVEKLSGDESS-----SETRGILIDINDPSTNNVNE 841
 Db 517 TESLTDNFVDVPEYKQPA-----IKKIFTDENIFQYLYSQTFP-LDIRDISUTSSPDD 570
 Qy 842 VHDASNTQGSVNTSDITNGHSESSLNRTTNAQDIKIGRSG-----NEQSDN-----QENS 892
 Db 571 ALLFSNKKVYFFSMDYI-----KTAN-----KVVEAGLPAGWVKQIVNDFVIEANK 616
 Qy 893 SHSDNSG-----SLTIGQVPESEDNTONTYDSN-----PHRDTNALASL 933
 Db 617 SNTMDKIADISLIVPYIGLALNVGNETAKGNFENAFIAGASILLLEFIPELLIPVVGAF 676
 Qy 934 PS---DDKINEIRGDSRDSRSGRGD-----TTSNT--HDOVR----- 966
 Db 677 LESYIDNKKIITIDNALTNRNEKSDMTGLIVAQWLSVTNTQFYTIKEGMYKALNYQA 736
 Qy 967 -----RTNIVSERRVNSHDFIRNGMAN--NNAHQYITQIEN--NG----- 1003
 Db 737 QALEEIIKYRYNIYSEKESKNINIDFNDINSKLNENGINAIDNINNFINGCSVSYLMKCM 796
 Qy 1004 -----IIRGQESAGNSVNYKDNPKRNSPSSENDHKKNIQEYNSR-----DTKRVREI 1052
 Db 797 IPLAVEKLLDFDNTLKKNLNLYIDENKLYLIGSAEYKSKVNYLAKTIMPPDLSIYNTD 856
 Qy 1053 ILSKQKCNKN 1063
 Db 857 ILEMFNKYN 867

RESULT 14

US-11-077-550-22
 ; Sequence 22, Application US/11077550
 ; Publication No. US20050244435A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Shone, Clifford Charles
 ; APPLICANT: Quinn, Conrad Padraig
 ; APPLICANT: Foster, Keith Alan
 ; APPLICANT: Chaddock, John
 ; APPLICANT: Marks, Philip
 ; APPLICANT: Sutton, J. Mark
 ; APPLICANT: Stancombe, Patrick
 ; APPLICANT: Wayne, Jonathan
 ; TITLE OF INVENTION: Recombinant Toxin Fragments
 ; FILE REFERENCE: 1581.0130004
 ; CURRENT APPLICATION NUMBER: US/11/077,550
 ; CURRENT FILING DATE: 2005-03-11
 ; PRIOR FILING DATE: 10/241,596
 ; PRIOR FILING DATE: 2002-09-12
 ; PRIOR FILING DATE: 09/255,829
 ; PRIOR FILING DATE: 1999-02-23
 ; PRIOR FILING DATE: PCT/GB97/02273
 ; PRIOR FILING DATE: 1997-08-22
 ; PRIOR FILING DATE: 08/782,893

Qy 967 -----RTNIVSERRVNSHDFIRNGMAN--NNAHHQYITQIEN--NG----- 1003
Db 740 ALNYQAALKEEIIKYRYNIYSEKESKININIDFNDSKLNKNEGAINNNFNGCSVS 799
Qy 1004 -----IIRGOESAGSNVYKNDPKESNFSSSENDHKKNIQYNSR-----DTK 1046
Db 800 YLMKKMIPLAVEKLLDFDNTLKKLLNYIDENKLYLIGSABEYKSKVNYKLTIMPFDLS 859
Qy 1047 RVREIIKLSKONKCNN 1063
Db 860 IYTNDTILIEMFNKYS 876

RESULT 17
US-11-077-550-106
; Sequence 106, Application US/11077550
; Publication No. US2005024435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 106
; LENGTH: 876
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-11-077-550-106

Query Match 2.1%; Score 136.5; DB 7; Length 876;
Best Local Similarity 18.5%; Pred. No. 0.066;
Matches 181; Conservative 139; Mismatches 346; Indels 311; Gaps 45;

Qy 252 TDRVKGINTKFSYYKEKVEKLNITKK-----BWEKKNANLWNVHVKGNISKECA 307
Db 46 TDRI-WIPERYTFGYPKDFNKSNGIFNRDVCYYDPYLN-----TNDKKN----- 93
Qy 308 IIPAEPPQINLWIKENWENFLMEKKRLPLAIKKCKVENKKEACFGCC-----RLPCSS 361
Db 94 -----FLOTWIKLFNRKSKPLGKLEMLIINGIPYLGORRVPLEE 134
Qy 362 YTFPMKSKTQMEVLTNLYKKNSGVDKNNFLNDFPKNNKNDLDDFFKNEKEYDDLCDC 421
Db 135 FN-----TNI-----ASVTVNKLISNPGVERKGI-----FAN----- 163
Qy 422 RYTATIISFLNGPAKD-----VDIASQINVDNLRGFG-----C-NYKS--NNEKSWNC 468
Db 164 -----LIIFGPGVLTNENETIDIGIQNHFASSREGFGGIMQKPCPEYVSFVNNVQ----- 214
Qy 469 TGTFTNKPFGTCEPRRQTLCLGRTYLLHRGHEEDYKEHLLGASIVEAQLLYKYEKDE 528

Db 215 -----NK-----GASIFNRRGYFSDPALILMHLELIHVLHGL-YGIKVDD- 252
Qy 529 NALCSIIIONS---YADLADIIGSDII-----KDYTYGKKME-----ENL 564
Db 253 ---LPVNPKEKFFMQSTDAIQAEELYTFGGQDPSIITPSTDKSIYDKVLQNRGIVDRL 309
Qy 565 NKVNDKKRNEBSLKIIFREKWD-----ENKENVWKMGAVALKXKTKCDYDFQKIPQF 619
Db 310 NKVLVCISDPNININITYKQKFKDKYKFVEDSEKYSI-----DVESFDKLYKS 357
Qy 620 LRW-FKEW--GDDFCERKEKIIYSPESFKVECKKDCDCENTCKNCKSEYKKWIDLKSEY 676
Db 358 LMFGFTETNIAENYKIKTRASYFSDSLPVKIK-----NLLDNEIYTIIEGFIKSDKM 411
Qy 677 EKQVDKYTKDKKKMYDNID-----EVKNKEANVYLKE-----KSECKEKNVDFD 721
Db 412 EKEVYRQNKAINKQAVEEISKEHLAVYKIQMCKSEKLYDDDDKDRWSSRICIDVDNED 471
Qy 722 KIPNESPNEYEDMCKKDEIKYLYNEIKYPTKHDIYDIDTFSTFGGTPGISANINEQ 781
Db 472 LFFIADKNSFSDLSKNERIEYNTQSNYIENDPFINELILDTDL-----ISKI 519
Qy 782 QSGKDTSGTNGSETSDSPVSHPEPESDAAINVEKLSGDESS-----SETRGILDINDPSV 835
Db 520 ELPSENTESLTDNVDVPTKQPA-----IKKIFTDENTIFYOYLYSQTPP-LDIRDISL 573
Qy 836 TNNVNEVHDASNTQGSVNTSDITNGHSSSLNRTTNAQDIKIGRSG-----NEQSDN-- 888
Db 574 TSSFDDALLFSNKVYFSDYI-----KTAN-----KVEAGLFAGWVKQIVNDF 619
Qy 889 --QENSSHSDNSG-----SLTIGQVPSDONTQNTYDSQN-----PHRDTP 927
Db 620 VIEANKSTMDKIADISLIVPYTIGLALNVGNETAKGNFENAFIAGASILLEFIPELLIP 679
Qy 928 NALASLPS--DDKINEIGFSDSRDSENGRGD-----TTSNT--HDVR----- 966
Db 680 VVGAFLLESYIDNKIKIITIDNALTNRNEKWSMDYGLIVAQLSTVNTQFYTIKEGMYK 739
Qy 967 -----RTNIVSERRVNSHDFIRNGMAN--NNAHHQYITQIEN--NG----- 1003
Db 740 ALNYQAALKEEIIKYRYNIYSEKESKININIDFNDSKLNKNEGAINNNFNGCSVS 799
Qy 1004 -----IIRGOESAGSNVYKNDPKESNFSSSENDHKKNIQYNSR-----DTK 1046
Db 800 YLMKKMIPLAVEKLLDFDNTLKKLLNYIDENKLYLIGSABEYKSKVNYKLTIMPFDLS 859
Qy 1047 RVREIIKLSKONKCNN 1063
Db 860 IYTNDTILIEMFNKYS 876

RESULT 18

US-11-077-550-108
; Sequence 108, Application US/11077550
; Publication No. US2005024435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273


```
Db 215 ----NK-----GASIFNRGYSFSDPALILMHLEHIVHGL-YGIVKDD- 252
Qy 529 NALCSIIIONS---YADLADIIGSDII-----KDYIGKOME-----ENL 564
Db 253 ---LPIVNEKFFMQSTDAIQABELYTFGQDPSIITPSTDKSIYDKVLQNRGIVDRL 309
Qy 565 NKYNKDKRNEESLKIFREKWD-----ENKENVWVMSAVLKNKTKDYDFKQKIPQF 619
Db 310 NKVLVCISDPNININIKYKPKDYKVFEDSEKYSI-----DVESFDKLYKS 357
Qy 620 LRW-FKEW--GDDFCERKEKIYFESFKVECKKDCDENTCKNKCSEYKKWIDLKSEY 676
Db 358 LMFGETTNAENYKIKTRASYFSDSLPPVKIK-----NLLDNEIYTIIEGFIISDKDM 411
Qy 677 EKQVDKYTKDKNKQWYDNIDVKNKEANVYLKESKE-----CKDVNFDKIFNESP 728
Db 412 EKEYRGONKAINQAY---EISKEHLAVYKIQMCSAIEGRAPGICIDVDNEDLFFIADK 468
Qy 729 NEYEDMCKKDEIKYLNEIKYKPKTHDIYDIDFSDTFGDTPIISINANINEQSGKDT 788
Db 469 NSPDDLSKNERIEYNTQSNYIENDFPINELILDIDL-----ISKIELPRENT 516
Qy 789 NTGNSSETSDSPVSHPEPSDAAINVEKLSGDSESS-----SETRGILDINDPSVTNNVNEV 842
Db 517 ESLTDFNVDVPVYEKQPA-----IKKIFTDENTIFQYLSQTFP-LDIRDISLTSFDDA 570
Qy 843 HDASNTQGSVNTSDITNGHSESSLNRTTNAQDIKIGRSG-----NQSDN-----QENSS 893
Db 571 LLSNKVYVFFSMDYI-----KTAN-----KVVEAGLFAGWVKQIVNDFVIEANKS 616
Qy 894 HSSDMSG-----SLTIGQVPSEDNTQNTYDSQN-----PHRDTNALASLP 934
Db 617 NTWMDKIADISLIVPYIGLALNVGNETAAGNFENAFIAGASILLEFTEPELLIPVVGAFLL 676
Qy 935 S---DDKINEIEGFDSSRDSSENGRGD-----TTSNT--HDVR----- 966
Db 677 ESYIDNKNKIKTIDNALTRENEKWSMDYGLIVAQWLSTVNTQPTTIKEGMYKALNYAQ 736
Qy 967 -----RTNIVSERVNSHDFIRNGMAN--NNAHHQYITQIEN--NG----- 1003
Db 737 ALBEEIKYRNIYSEKESKINIDFNDINSKLENGINQAINNFINFGCSVSYLMKKMI 796
Qy 1004 -----IIRQOESAGNSVNYKDKPKRSNFSSENHDKNQIENSR-----DTRKVRBEII 1053
Db 797 PLAVEKLLDFDNTLKNLNYIDENKLYLIGSAEYKSKVNYKYLKTIIMPFDLSIYNTDTI 856
Qy 1054 KLSKQKCNN 1063
Db 857 LIEMFNKYS 866
```

```
RESULT 20
US-11-077-550-175
; Sequence 175, Application US/11077550
; Publication No. US2005024435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/255,829
; PRIOR FILING DATE: 1999-02-23
```

```
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 175
; LENGTH: 860
; TYPE: PRT
; ORGANISM: Clostridium botulinum
; US-11-077-550-175

Query Match 2.1%; Score 133.5; DB 7; Length 860;
Best Local Similarity 18.6%; Pred. No. 0.097;
Matches 180; Conservative 137; Mismatches 346; Indels 307; Gaps 45;

Qy 252 TDRVGYINTKFSYDYKRNVEKLNKIKK-----EWMKKNANLWNHMIWNHKNISKECA 307
Db 40 TDRI-WIIPERYTFGYKPEDFNKSGIFNRDVCEYDPPVYN-----TNDKKNI----- 87
Qy 308 IIPAEFQINLWIKENWENFLMEKKELFLNLIKDKCVENKYEACFGC-----RLPSS 361
Db 88 -----FLQTMIKLFNRIKSKPLGKLELMIINGIPYLGRRVRPLEE 128
Qy 362 YTSFMKSKTQMEVLNLYKKNQSGVDKNNFLNDLFLKNNKNDLDDFFKNEKEYDLDLDC 421
Db 129 FN-----TNI-----ASVTNKLISNPGEVERKGI-----FAN----- 157
Qy 422 RYTATIIKSFLNGPAKND---VDIASQINVDNLRGFG-----C-NYKS--NNEKSWNC 468
Db 158 -----LIIFGCPVLNENETIDIGIQNHPASREGFGGIMQMKFCPEYVSFVFNQV 208
Qy 469 TGTFTNKFPGTCBPPRRQTLCLGRTYLLHRGHEEDYKEHLGLGASIEAQLLYKYEKDE 528
Db 209 -----NK-----GASIFNRGYSFSDPALILMHLEHIVHGL-YGIVKDD- 246
Qy 529 NALCSIIIONS---YADLADIIGSDII-----KDYIGKOME-----ENL 564
Db 247 ---LPIVNEKFFMQSTDAIQABELYTFGQDPSIITPSTDKSIYDKVLQNRGIVDRL 303
Qy 565 NKYNKDKRNEESLKIFREKWD-----ENKENVWVMSAVLKNKTKDYDFKQKIPQF 619
Db 304 NKVLVCISDPNININIKYKPKDYKVFEDSEKYSI-----DVESFDKLYKS 351
Qy 620 LRW-FKEW--GDDFCERKEKIYFESFKVECKKDCDENTCKNKCSEYKKWIDLKSEY 676
Db 352 LMFGETTNAENYKIKTRASYFSDSLPPVKIK-----NLLDNEIYTIIEGFIISDKDM 405
Qy 677 EKQVDKYTKDKNKQWYDNIDVKNKEANVYLKESKE-----CKDVNFDKIFNESP 728
Db 406 EKEYRGONKAINQAY---EISKEHLAVYKIQMCSAIEGRAPGICIDVDNEDLFFIADK 462
Qy 729 NEYEDMCKKDEIKYLNEIKYKPKTHDIYDIDFSDTFGDTPIISINANINEQSGKDT 788
Db 463 NSPDDLSKNERIEYNTQSNYIENDFPINELILDIDL-----ISKIELPRENT 510
Qy 789 NTGNSSETSDSPVSHPEPSDAAINVEKLSGDSESS-----SETRGILDINDPSVTNNVNEV 842
Db 511 ESLTDFNVDVPVYEKQPA-----IKKIFTDENTIFQYLSQTFP-LDIRDISLTSFDDA 564
Qy 843 HDASNTQGSVNTSDITNGHSESSLNRTTNAQDIKIGRSG-----NQSDN-----QENSS 893
Db 565 LLSNKVYVFFSMDYI-----KTAN-----KVVEAGLFAGWVKQIVNDFVIEANKS 610
Qy 894 HSSDMSG-----SLTIGQVPSEDNTQNTYDSQN-----PHRDTNALASLP 934
Db 611 NTWMDKIADISLIVPYIGLALNVGNETAAGNFENAFIAGASILLEFTEPELLIPVVGAFLL 670
Qy 935 S---DDKINEIEGFDSSRDSSENGRGD-----TTSNT--HDVR----- 966
```

Db 671 ESYDINKNKKIITDNLTKRNEKWSMDYGLIQAQWLSVTNTQFYTIKBMGYKALNYQAQ 730
 Qy 967 -----RTNIVSERVNSHDFIRNGMAN--NNAHQYITQIEN--NG-----1003
 Db 731 ALBEIIKRYNIYSEKESNINIDFNDINSKLNKNGINQAINNFINFGSCSVYLMKMI 790
 Qy 1004 -----IIRQESAGNSVNYKPNKRSNFSNDHKHKNIOEYNSR-----DTRKVRBEII 1053
 Db 791 PLAVEKLLDPNTLKNLLNIDENKLYLIGSABEYKSVKNYKLTIMPFDLSIYTNDTI 850
 Qy 1054 KLSQKNCNN 1063
 Db 851 LIEMFNKYS 860

RESULT 21

US-11-077-550-94
 ; Sequence 94, Application US/11077550
 ; Publication No. US20050244435A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Shone, Clifford Charles
 ; APPLICANT: Quinn, Conrad Padraig
 ; APPLICANT: Foster, Keith Alan
 ; APPLICANT: Chaddock, John
 ; APPLICANT: Marks, Philip
 ; APPLICANT: Sutton, J. Mark
 ; APPLICANT: Stancombe, Patrick
 ; APPLICANT: Wayne, Jonathan
 ; TITLE OF INVENTION: Recombinant Toxin Fragments
 ; FILE REFERENCE: 1581.0130004
 ; CURRENT APPLICATION NUMBER: US/11/077,550
 ; PRIOR FILING DATE: 2005-03-11
 ; PRIOR FILING DATE: 2002-09-12
 ; PRIOR APPLICATION NUMBER: 09/255,829
 ; PRIOR FILING DATE: 1999-02-23
 ; PRIOR APPLICATION NUMBER: PCT/GB97/02273
 ; PRIOR FILING DATE: 1997-08-22
 ; PRIOR APPLICATION NUMBER: 08/782,893
 ; PRIOR FILING DATE: 1996-12-27
 ; PRIOR APPLICATION NUMBER: GB9625996.5
 ; PRIOR FILING DATE: 1996-12-13
 ; PRIOR APPLICATION NUMBER: GB9617671.4
 ; PRIOR FILING DATE: 1996-08-23
 ; NUMBER OF SEQ ID NOS: 179
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 94
 ; LENGTH: 862
 ; TYPE: PRT
 ; ORGANISM: Clostridium botulinum
 US-11-077-550-94

Query Match 2.1%; Score 133.5; DB 7; Length 862;
 Best Local Similarity 18.6%; Pred. No. 0.097;
 Matches 180; Conservative 137; Mismatches 346; Indels 307; Gaps 45;
 Qy 252 TDRVKGINTFSDYKKEKVEKLNKIKK-----EWEKKNALNWHVHVNKHGNIKECA 307
 Db 42 TDRI-WIIPERYTFGYKPEDFNKSSGIFNRDVCEYDPOYLN-----TNDKQNI-----89
 Qy 308 IIPABEPQINLWIKENWFLEWKRLFLNLIKDKOVENKYEACFGGC-----RLPCS 361
 Db 90 -----FLOTWIKLFNRKSLPGLKEKLEMIINGIPYLGDRVPLEE 130
 Qy 362 YTFPMKSKTQMBVLTNLYKKNSGVDKNNFLNDFKNNKNDLDPFKNEKEDDLDC 421
 Db 131 FN-----TNI-----ASVTNKLISNPGEVERKGI-----PAN-----159
 Qy 422 RYTATIISFLNGPAKD-----VDIASQINVDNLRGFG-----C-NYKS--NNEKSWNC 468
 Db 160 -----LIIFGPGVLENETIDIGIHNHFASREGFGGIMQKFCPEYVSFVNNVQE---210

Qy 469 TGFTNKFPGTCPPRRQTILGRYLLHRHGHEBDYKEHLIGASIVYBAQLIKYKKEKDE 528
 Db 211 -----NK-----GASIFNRRGYFSDPALILMHHLIHLVHGL-YGIKVDD-248
 Qy 529 NALCSIIIONS---YADLADIKGSII-----KDYKQKME-----ENL 564
 Db 249 ---LPIVPNEKCFMQSTDAIQABELYTFGQDPSIIITPSTDKSIYDKVLQNFGRGIVDRL 305
 Qy 565 NKVNDKRRNEESLKIIFREKWD-----ENKENVWVMSAVLKNKCTCKDYDKFQKIQOF 619
 Db 306 NKVLVCLISDPNINININYKAFKDYKVFDESEKYSI-----DVESEKLYKS 353
 Qy 620 LRW-FKEW--GDDPCEKREKIYSPFSKVECKKDCDENTCKNKCSEYKWIIDLKSEY 676
 Db 354 LMFQFTETNAENYKIKTRASYFSDSLPPVKIK-----NLLDNEIYTBEGFNISDKDM 407
 Qy 677 EKQVDKYTKKKNKMYNDIDEVKNKEANVYLKESKE-----CKDVNPDDEKIFNESP 728
 Db 408 EKEYRGQNKAINKOAY---EEISKEHLAVYKIQMCISAIEGRAPGICIDVDNEDLFFIADK 464
 Qy 729 NEYEDMCKCDEIKYLNKIKYPTKTHIDYDIDTFSDTFGDTPTISINANINEQSQKDT 788
 Db 465 NSFSDLSKNERIEYNTQSNYIENDFPINELILDIDL-----ISKILPSENT 512
 Qy 789 NTGNSSETSDSPVSHESPESDAAINVEKLSGDESS-----SETRGILDINDPSVTNNVNEV 842
 Db 513 ESLTDFENVDPVVEKQPA-----IKKIFTDENTIFQYLYSQTTP-LDIRDISLTSSPDDA 566
 Qy 843 HDASNTQGSVNTSDITNGHSESLNRTTNAQDIKIGRSQ-----NQSDN-----QENS 893
 Db 567 LLFSNKVYSPFSDYI-----KTAN---KVEAGLIFAGWVKQIVNDFVIEANKS 612
 Qy 894 HSSDSNG-----SLTIGQVPSDNTONTYDSQ-----PHRDTNALASLP 934
 Db 613 NTMDKIADISLIVPIYIGLALNVGNETAGNFENAFETAGASILLLEFPELLIPVVGAPLL 672
 Qy 935 S---DDKINEIEGPDSSRSENGRGD-----TTSNT---HDVR-----966
 Db 673 ESYDINKNKKIITDNLTKRNEKWSMDYGLIQAQWLSVTNTQFYTIKBMGYKALNYQAQ 732
 Qy 967 -----RTNIVSERVNSHDFIRNGMAN--NNAHQYITQIEN--NG-----1003
 Db 733 ALBEIIKRYNIYSEKESNINIDFNDINSKLNKNGINQAINNFINFGSCSVYLMKMI 792
 Qy 1004 -----IIRQESAGNSVNYKPNKRSNFSNDHKHKNIOEYNSR-----DTRKVRBEII 1053
 Db 793 PLAVEKLLDPNTLKNLLNIDENKLYLIGSABEYKSVKNYKLTIMPFDLSIYTNDTI 852
 Qy 1054 KLSQKNCNN 1063
 Db 853 LIEMFNKYS 862

RESULT 22

US-11-077-550-88
 ; Sequence 88, Application US/11077550
 ; Publication No. US20050244435A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Shone, Clifford Charles
 ; APPLICANT: Quinn, Conrad Padraig
 ; APPLICANT: Foster, Keith Alan
 ; APPLICANT: Chaddock, John
 ; APPLICANT: Marks, Philip
 ; APPLICANT: Sutton, J. Mark
 ; APPLICANT: Stancombe, Patrick
 ; APPLICANT: Wayne, Jonathan
 ; TITLE OF INVENTION: Recombinant Toxin Fragments
 ; FILE REFERENCE: 1581.0130004
 ; CURRENT APPLICATION NUMBER: US/11/077,550
 ; CURRENT FILING DATE: 2005-03-11
 ; PRIOR FILING DATE: 2002-09-12
 ; PRIOR APPLICATION NUMBER: 09/255,829

Qy 469 TGTFTNKPCTGCPERRQTLCLGRYTLHLRHGHEEDYKEHLGASIEYEAQLLYKYKKEKDE 528
Db 216 ---NK-----GASIFNRKGFPSDPAIILMHHLHVLHGL-YGKVVDD- 253
Qy 529 NALCSIIQNS---YADLADIIGSDII-----KDYKKGME-----ENL 564
Db 254 ---LPVNEKPFQMSQTDALQAEELYTFGQDPSIITPSTDKSIYDKVLQNFPGIVDRL 310
Qy 565 NKVNDKCKRNEESLKIFREKWD-----ENKENVKVMSAVLKNKETCKDYKFKQKIQF 619
Db 311 NKVLVCISDPNININIKNFKDKYFVEDSEGKYSI-----DVESFDKLYKS 358
Qy 620 LRW-FKEW--GDDPCEKKEKIEYFESFKVECKKCCDENTCKNKCSEYKWKWIDLKSEY 676
Db 359 LMFGETETNAENYKIKTRASYFSDSLPPVKIK-----NLLDNEIYTIIEGFNISDKDM 412
Qy 677 EKQVDKTKDKNKKMYDNIDVENKKEANVYLKESKSE-----CKDVNFDKIFNESP 728
Db 413 EKEYRGONKAINQAY---EISKEHLAVYKIQMCSAIEGRAPGICIDVNDLFFIADK 469
Qy 729 NEYEDMCKCKDEIKYLNIEIKYPTKHDIDYDITFSDTFGDTGPISINANIEQSGKDT 788
Db 470 NSPDDLSKNERIEYNTQSNYIENDFPINELILDIDL-----ISKIELPSENT 517
Qy 789 NTGNSSETSDSPVSHPESDAANVEKLSGDSESS-----SETRGILDINDPSVTNNVNEV 842
Db 518 ESLSLTDNFVDPVVEKQPA-----IKKIFTDENTIFQVLYSQTFP-LDIRDISLTSSPDDA 571
Qy 843 HDASNTQGSVNSNTSDITNGHSESSLRNTAQAQIKIGRSG-----NEQSDN-----QENSS 893
Db 572 LLFSNKVYSPFSDYI-----KTAN-----KVVEAGLFAGVWKQIVNDFVIEANKS 617
Qy 894 HSSDNGS-----SLTIGQVPSDNTQNTYDSON-----PHRDTNALASLP 934
Db 618 NTMDKIADISLIVPYIGLALNVGNETAAGNFENAFETAGASILLEFPELLIPVVGAFLL 677
Qy 935 S---DDKINIEGDSRDSRDSNGRD-----TTSNT---HDVR-----966
Db 678 ESYIDNKKIKITIDNALTKENEKWSMDYGLIVAQWLSVTNTQFYTIKEGMYKALYQAQ 737
Qy 967 -----RNVISERVNSHDFIRNGMAN--NNAHQYITQIEN--NG-----1003
Db 738 ALBEEKRYNIYSEKSNINIDFNDSKLNENGINQAIDNINNFNGCSVSLMKQMI 797
Qy 1004 -----IIRQESAGNSVNYKNPKESNFSSENDHKKIQEYNSR-----DYKREVEII 1053
Db 798 PLAVEKLLDFNTLKKLLNLYDENKLYLGSAEYKSKVKNYKLTIMPFDLSIYTTNDTI 857
Qy 1054 KLSQKQKCNN 1063
Db 858 LIEMFNKYS 867

RESULT 24
US-11-077-550-98
; Sequence 98, Application US/11077550
; Publication No. US2005024435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12

; PRIOR APPLICATION NUMBER: 09/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625596.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 98
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-11-077-550-98

Query Match 2.1%, Score 133.5; DB 7; Length 867;
Best Local Similarity 18.6%; Pred. No. 0.098;
Matches 180; Conservative 137; Mismatches 346; Indels 307; Gaps 45;
Qy 252 TDRVKGIVNTKFSYDYKERNVEKLNINIKK-----EWEKKNKANLNMHVMVNHKGNISKECA 307
Db 47 TDRI-WIIPERYTFGYKPEDFNKSSGIFNRDVCYDYLN-----TNDKKNI-----94
Qy 308 IIPAEPEQINLWIKENENFLMEKRLFLNIDKCKVENKKEACFGGC-----RLPSS 361
Db 95 -----FLOTMIKLFNRIKSKPLGEKLEMLIINGIPYIGDRRVPLUEE 135
Qy 362 YTSFMKSKTKQMEVLTNLYKKQNSGVNDKNNFLNDLPKNNKNKNDLDDFFKNEKEYDDLDLDC 421
Db 136 FN-----TWI-----ASVTNKLISNPGVERKKGII-----FAN-----164
Qy 422 RYTATIKSFLNGPAKND---VDIASQINVNDLRGFG-----C-NYKS---NNEKSWNC 468
Db 165 -----LIIIFGPGVLNENETIDIGIQNHFPASREGFGGIMQMFKFCEYVSVFNNVQE---215
Qy 469 TGTFTNKPCTGCPERRQTLCLGRYTLHLRHGHEEDYKEHLGASIEYEAQLLYKYKKEKDE 528
Db 216 ---NK-----GASIFNRKGFPSDPAIILMHHLHVLHGL-YGKVVDD- 253
Qy 529 NALCSIIQNS---YADLADIIGSDII-----KDYKKGME-----ENL 564
Db 254 ---LPVNEKPFQMSQTDALQAEELYTFGQDPSIITPSTDKSIYDKVLQNFPGIVDRL 310
Qy 565 NKVNDKCKRNEESLKIFREKWD-----ENKENVKVMSAVLKNKETCKDYKFKQKIQF 619
Db 311 NKVLVCISDPNINININIKNFKDKYFVEDSEGKYSI-----DVESFDKLYKS 358
Qy 620 LRW-FKEW--GDDPCEKKEKIEYFESFKVECKKCCDENTCKNKCSEYKWKWIDLKSEY 676
Db 359 LMFGETETNAENYKIKTRASYFSDSLPPVKIK-----NLLDNEIYTIIEGFNISDKDM 412
Qy 677 EKQVDKTKDKNKKMYDNIDVENKKEANVYLKESKSE-----CKDVNFDKIFNESP 728
Db 413 EKEYRGONKAINQAY---EISKEHLAVYKIQMCSAIEGRAPGICIDVNDLFFIADK 469
Qy 729 NEYEDMCKCKDEIKYLNIEIKYPTKHDIDYDITFSDTFGDTGPISINANIEQSGKDT 788
Db 470 NSPDDLSKNERIEYNTQSNYIENDFPINELILDIDL-----ISKIELPSENT 517
Qy 789 NTGNSSETSDSPVSHPESDAANVEKLSGDSESS-----SETRGILDINDPSVTNNVNEV 842
Db 518 ESLSLTDNFVDPVVEKQPA-----IKKIFTDENTIFQVLYSQTFP-LDIRDISLTSSPDDA 571
Qy 843 HDASNTQGSVNSNTSDITNGHSESSLRNTAQAQIKIGRSG-----NEQSDN-----QENSS 893
Db 572 LLFSNKVYSPFSDYI-----KTAN-----KVVEAGLFAGVWKQIVNDFVIEANKS 617
Qy 894 HSSDNGS-----SLTIGQVPSDNTQNTYDSON-----PHRDTNALASLP 934
Db 618 NTMDKIADISLIVPYIGLALNVGNETAAGNFENAFETAGASILLEFPELLIPVVGAFLL 677

Db 622 NTMDKIADISLIVPIGLANVGNETAKGNFENAFIAGASILLPEIPPELLIPVVGAFLL 681
Qy 935 S---DDKINEIEGDSRDSNGRGD-----TTSNT--HDVR----- 966
Db 682 ESYIDNKKIITIDNALTKRNEKWSMDYGLIVAQWLSTVNTQYTIKEGMYKALNYQAQ 741
Qy 967 -----RTNIVSERVNSHDFIRNGMAN--NNAHQYITQIEN--NG----- 1003
Db 742 ALBEEIKYRYNIYEKEKSNINIDPNDINSKLNIGINQAIDNINNFINGCSVSVMKKMI 801
Qy 1004 -----IIRGOESAGNSVNYKDPKPSNFSSNDHKNIQIENSR-----DTRKRVREII 1053
Db 802 PLAVEKLLDFNTLKKLLNLYIDENKLYLIGSAEYKSKVKNYKLTIMPFDLSIYTNDTI 861
Qy 1054 KLSKQNKCN 1063
Db 862 LIEMFNKYS 871

RESULT 29

US-11-077-550-171
; Sequence 171, Application US/11077550
; Publication No. US20050244435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stacombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; PRIOR FILING DATE: 2005-03-11
; PRIOR FILING DATE: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 03/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 171
; LENGTH: 862
; TYPE: PRP
; ORGANISM: Clostridium botulinum
US-11-077-550-171

Query Match 2.0%; Score 132.5; DB 7; Length 862;

Best Local Similarity 18.5%; Pred. No. 0.11;

Matches 180; Conservative 137; Mismatches 346; Indels 309; Gaps 45;

Qy 252 TDRVKGINTKFSYDYKKNVKNLNNIKK-----EWEKKNKALNWHMIVNKHGNIKECA 307
Db 40 TDRI-WIIPERYTFGYPKDPNKGSGIFNRDVCYIYDPDYLN-----TNDKKN----- 87
Qy 308 IIPAEFPQINLWIKENWFNMEKRLFLNFKDCVENKRYKACFGG-----RLPCSS 361
Db 88 -----FLQTWIKLFNRKSKPLGKLEMLINGIPYIGDRVRPLEE 128
Qy 362 YTSFPMKSKTQMEVLTNLYKKNSGVDRKNPFLNDFKKNQNDLDDFFKNEKEYDDLDC 421
Db 129 FN-----TNI-----ASVTNKLISNPGVERKKG-----FAN----- 157

Qy 422 RYTATIKSLPNGPAKND---VDIASQINVNDLRGFG-----C-NYKS--NNEKSWNC 468
Db 158 -----LIIFGPGVLNENETIDIGIQNHFPASREGGIMQMKFCPPYVSFVNNVQ-- 208
Qy 469 TGTFTNKPFCPTCBPPRQTLCLGRYLLHHRGHBEDYKHLGLGASIIYBAQLLKYYKKEKDE 528
Db 209 -----NK-----GASIFNRGRYSDPALILMHILHVLHGL-YGIKVDD- 246
Qy 529 NALCSIIQNS---YADLADIIGSDII-----KDYCKKKE-----ENL 564
Db 247 ---LPIVPNEKKFPMGSTDAIQABELYTFGQDPSIITPSTDKSIYDKVQNFRGIVDRL 303
Qy 565 NKYNKDKKRNESEIKIFREKWD-----ENKENVMKMSAVLKNKTKCDYDFKQKIPQF 619
Db 304 NKVLICISDPNININIKKFKDKYKFEVDESEKYSI-----DVESFDKLYKS 351
Qy 620 LRW-FKEW--GDDPCEKREKIIYFSESFKVECKKCCDENTCKNKCEYKKNWIDLKSEY 676
Db 352 LMFGFTETNIAENYKIKTRASYFSDSLPPVKIK-----NLLDNEIYITIBEGFNISDKDM 405
Qy 677 EKQVDKYTKDKNKKMYDNIDEVKNKEANVYLKEKSE-----CKDWNFDKIFNE 726
Db 406 EKYRGONKAINQAY---EESKEHLAVYKIQMCSAIEGRCSAPGICIDVDNEDLFIA 462
Qy 727 SPNEYEDMCKKDBIKYLNIEIKYPTKHDYIDITFSDTFDGTGPISINANINEQQSGKD 786
Db 463 DKNSPDDLSKNERIEVNTQSNIENDFPINELLIDTDL-----ISKIELPSE 510
Qy 787 TSNTGNSETSDSPVSHPEPSDAAINVEKLSGDESS-----SETRGILDINDPSVTNNVN 840
Db 511 NTESLTDNFNDVPVYEKQPA-----IKKIPTDENTIFQYLYSQTFF-LDIRDISLTSFSD 564
Qy 841 EVHDAANTQGSVNTSDITNGHSESSLNRTNAQDIKIGSG-----NEQSDN-----QEN 891
Db 565 DALLFSNKVYSPFMDYI-----KTAN-----KVEAGLFGAGVWKQIVNDFVIEAN 610
Qy 892 SSHSSDNGS-----SLATIGQVPSDNTQNTYDSQN-----PHRDTPNALAS 932
Db 611 KSNMTDKIADISLIVPIGLANVGNETAKGNFENAFIAGASILLPEIPPELLIPVVGAF 670
Qy 933 LPS---DDKINEIEGDSRDSNGRGD-----TTSNT--HDVR----- 966
Db 671 LLESYIDNKKIITIDNALTKRNEKWSMDYGLIVAQWLSTVNTQYTIKEGMYKALNYQ 730
Qy 967 -----RTNIVSERVNSHDFIRNGMAN--NNAHQYITQIEN--NG----- 1003
Db 731 AQALEEIKYRYNIYEKEKSNINIDPNDINSKLNIGINQAIDNINNFINGCSVSVMKK 790
Qy 1004 -----IIRGOESAGNSVNYKDPKPSNFSSNDHKNIQIENSR-----DTRKRVRE 1051
Db 791 MIPLAVEKLLDFNTLKKLLNLYIDENKLYLIGSAEYKSKVKNYKLTIMPFDLSIYTND 850
Qy 1052 IIKLSKQNKCN 1063
Db 851 LIEMFNKYS 862

RESULT 30

US-11-077-550-167
; Sequence 167, Application US/11077550
; Publication No. US20050244435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stacombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550

```
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 167
; LENGTH: 873
; TYPE: PRT
; ORGANISM: Clostridium botulinum
; US-11-077-550-167

Query Match          2.0%; Score 130.5; DB 7; Length 873;
Best Local Similarity 19.2%; Pred. No. 0.15;
Matches 163; Conservative 123; Mismatches 322; Indels 239; Gaps 40;

QY 388 DKWFL---NDLFKKNKNDLDDFFKNE-----KEYDLC-----DC 421
Db 85 EKONYLKGVTKFLERIYSTDLGRMLTTSIVRGIPFMGGSTIDTELKVIDTNCINVIQPDG 144
QY 422 RYTATIIKSFLNPAKNDVDIASQINVDNLGRGFCNYKSNKSNWCT-----GTFNKK 475
Db 145 SYRSEELNVILGSDIIOF-----ECKSGHEVLNTRNGYSGTQVIRFSPDFTFG 197
QY 476 FPGTCEPRRQTLCLGR-----TYLLHRGHEEDYKEHLGLGASI-----513
Db 198 FEESLEVDWTPLLGAGKFPATPAVLAHQLIHAG-----RLYGAINENRVFKYNTNA 251
QY 514 -YEAQLKYYKYE-----KDNALCSIIQNSYADLADII-KGSDII-----552
Db 252 YYEMSGLEVFELRTFGGHDAKFIDSLQENEFRLYYNKFQDIASLTKAKSIVGTTAS 311
QY 553 ----KDYGKK---MEENLNKVNKKRNEESLKIPEKRWMDENKENVKVMKMSAVLKNKE 605
Db 312 LQWKVNFKEKYLSEBTSKFSVDLKLKPKLYKMLTEIYEDNFVKFPVL-----NRK 366
QY 606 TKDYD----KFQKIPQ-----FLRFKEWGGDDFCERKEKIYSPE 642
Db 367 TYLNFDAKAVFKINIVPKVNYTIYDGFNLRLNLAANFNGQNTNINNNFTKLNFTGLPE 426
QY 643 SFKVEC-----KKKDCDENTCK---NKCSEYKKWIDLKKSBEYKQVDKYTKDKKNQMY 692
Db 427 FYKLLCVRGIIITSKTSLDRGYNKALMDLCIKVNNW-DLFFSPSE---DNFTNDLNKGE 482
QY 693 DNID-EVKNKEANVYLKEKSECKDVNFDKIFNES-PNEYEDMKCKDCBIKYLNEIKYP 750
Db 483 ITSDTNIEAAENISLDLIOQYLTFTFNDEPENISLENSSDIIIGOLEMP-NIERFP 540
QY 751 KTKHDIYDIDTFSDTGDGTPPISINANINQQSGKTSNTGNSSETSDSPVSHPEPSDAAI 810
Db 541 NGKK--YELDKYT-----MFHYLRAQFEHGKSRIALTNS-----VNEALL 579
QY 811 NVEKLSGDESSSTRGLINDPVSNTNNVNEVHDASNTQGSVN-TSDITNGHSESSLNR 869
Db 580 NPSRVYTFSSDYVK-----KVNKATEAAMFLGWVQLVYDFDTSE--VST 625
QY 870 TTNAQDIKIGRSGNEQSDNQSNGSGSLTIG-----OVPSDNTQNTY 917
Db 626 TDKIADITIIPIYGPAINLGNMLYKDDFVGALIFSGAVILLBFIPIAIP-----VLGTF 681
QY 918 DSONPHRDTNALASLPDSDDKINEIEGFDSSRDSNGRGDTTNTNTHDVRTNIVSERRVN 977
Db 682 ----ALVSVIA-NKVLTVQTDIDNALSKRNEKWD---EVYKYIVTNWLA--KVN 724
```

```
978 SH-DFIRNGWANNNAHQVITOENNGIIRGOEESAGNSVNYKDNPKRNSFFSENDRHKN 1036
725 TQIDLIRKKMKE-----ALEN-----QAEATKAIINYQN--QYTEEBKNNINFN 767
1037 IQEYNSRDTKRVREELIKLSK-QNKNNEYSMEYCYTYSBERNSPGPCSEERKKLCCOI 1095
768 IDLSKLNESINKAMININKFLNQCVSYLM-----NSMIPYGVKRLDFDASL 817
1096 SDYCLKY 1102
818 KDALLKY 824

RESULT 31
US-11-074-176-152
; Sequence 152, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Klaenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McAuliffe, Olivia
; APPLICANT: Peril, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore
; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; PRIOR FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 152
; LENGTH: 438
; TYPE: PRT
; ORGANISM: Lactobacillus acidophilus
; US-11-074-176-152

Query Match          2.0%; Score 130; DB 7; Length 438;
Best Local Similarity 19.5%; Pred. No. 0.072;
Matches 90; Conservative 72; Mismatches 157; Indels 142; Gaps 19;

QY 560 MEENLNKVNKKRNEESLKIPEKRWMDENKENVKVMKMSAVLKNKCTKDYDKFQKIPQF 619
Db 19 MNSTIVHADKGTSHBESSKVTKNDD--KNVPESEQETSSNNIDQSQRKEK---72
QY 620 LRWFKEWGGDDFCERKEKIYSPEKVECKKDCDENTCK---NKCSEYKKWIDLKSEY 676
Db 73 -----EQAIPEQDQSQNTNNDQNDASEDEDEDEVSVEDY 110
QY 677 EKQVDKYTKDKKNQMYDNIDEVKNKEANVYL-----KEKSECKD-----V 717
Db 111 ENNVKDFHVKQVQVKVQKLLAEKNNQHELMYIGRPTCYCQFSPDLKDFNEIVKGKLLYP 170
QY 718 NFDDKLFNESPNEYE-----DMCKKDCDEIKYLE-----IKYPTKHDIYDIDTFSDTFG 767
Db 171 NIDDE---EGAHDFAFKVIQIPGTPPTTMRPMNGKLISAWIGCKEKTQGLHDF-LYSDT-- 224
QY 768 DGTPTISINAN-INBQQSGKDTSGTNSSETSDSPVSHPEPSDAAINVEKLSGDESSSTRG 826
Db 225 -----ANKLVEQVVIKQNSNDTATQADNDVVASSEKTPTEVTVEENQAQSN----- 271
QY 827 ILDINDPVSNTNNV-NEV-HDASNTQGSVNTSDITNGHSE-----SSLNRTT----- 871
Db 272 ----NDVAITNFAENSVPENAKNVASSTADLTQVATGDQDDVAPKAETKNTKTVKPKIKHK 327
QY 872 -----NAQDIKIGRSGNEQSDNQSNGSGSLTIGQVSEEDNTQNTYDS 919
Db 328 IVANKVKQAKLHKHTNIIIPMSAKKREDYKEN-----NOYDT 364
QY 920 QNPHRDTNALASLPDSDDKINEIEGFDSSRDSNGRGDTTTS 960
Db 365 VKVHGTSFPAI-----KQKQARI---TMLKELENDTSDTIS 397
```

RESULT 32
US-11-077-550-2
; Sequence 2, Application US/11077550
; Publication No. US2005024435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR FILING DATE: 2005-03-11
; PRIOR FILING DATE: 2002-09-12
; PRIOR FILING DATE: 1999-02-23
; PRIOR FILING DATE: 1997-08-22
; PRIOR FILING DATE: 1996-12-13
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 871
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-11-077-550-2
Query Match 2.0%; Score 129.5; DB 7; Length 871;
Best Local Similarity 19.2%; Pred. No. 0.17;
Matches 163; Conservative 123; Mismatches 322; Indels 239; Gaps 40;
Qy 388 DKNFL---NDLFKNNKNDLDDFFKNE-----KEYDDL-----DC 421
Db 83 EKONYLKGVTKLFRYIYSDLGRMLTSIVRGIPFWGGSTIDELKVIDTNCINVIQPDG 142
Qy 422 RYATITIKSFLNGPAKNDVDIASQINVDNLRGFCNKYNNKSWNCT-----GTFTNK 475
Db 143 SYRSEELNLVIIGPSADIQF-----ECKSFGHEVLNLRNGYGSTQYIRSPDFTFG 195
Qy 476 PPGTCBPFRQTLCLGR-----TYLHRGHEEDYKXHLGASIT-----513
Db 196 FEESLEVDNTPLLGAGKFAFDPAVTLAHELHAGH-----RLYGIAINPRVFKVNTNA 249
Qy 514 -YEAQLKYYKKE-----KDNALCSIIQNSVADLADII-KGSDII-----552
Db 250 YEMSGLEVSFEELRTFGGHDAKAFIDSLQENEFRLYYNFKPOTASTLNKAKSVIGTAS 309
Qy 553 ----KDYVGK---MEENLNKVNKKRNEESLKFREKWDENKNNVKNVSAVLKNKE 605
Db 310 LOYMNKVPKYLISEDSGKFSVDKLFKLYKMLTEIYTFEDNFVFKVL-----NRK 364
Qy 606 TKDQYD----KFQKIPQ-----FLRNFKEWDDFCEKKEKIYSPE 642
Db 365 TYLNFDAVKFKNIVPKVNYTYIDGFLNRLNTLAANFNGQNTNMMNFTKLNFTGLFE 424
Qy 643 SPKVEK-----KKQCDENTCK---NKCSEYKWKIDLLKSEYKQVDKTKDKNNKY 692
Db 425 FYKLLCVRGIIITSKSLDKGYNKALNDLCIKVNNW-DLFFSPSE---DNFTDLNKGEE 480
Qy 693 DNID-EVKNKEANVYLKESKECKDVNFDKIFNES-PNEYEDMKCKDEIKYLNIEIKYP 750

Db 481 ITSNTNIEAAENISLDLIQYYLTFNFDNPEPENISLENLSSDIIGOLELMP--NIERRFP 538
Qy 751 KTKHDIYDITFSDTFGDTGTPISINANINEQOQKDTSTNCGSETSDSPVSHPESDAAI 810
Db 539 NGKK--YELDKY-----MPLYLRAQEFHGKSRIALTNS-----VNALL 577
Qy 811 NVEKLSDGESSESTRGILDINDPSVTNNVNEVHDASNTQGSVN--TSDITNGHSESSLNR 869
Db 578 NPSRVYTFSSDYVK-----KVNKATEAAMFLGWQLVYDFDDETSE--VST 623
Qy 870 TTNAQDIKIGRSGNEQSDNOENSHSDNSGSLTIG-----QVPSEDNTONTY 917
Db 624 TDKIADITIIPIYGALNIGNMLYKDDDFGALIFSGAVILLEPIPIAIP-----VLGTG 679
Qy 918 DSQNPREDTPNALASLPDDKINEIEGFSRDSRDSNGRGTTSNTHDVRTNIVSERRVN 977
Db 680 -----ALVSYIA-NKVLTVQIIDNALSKRNEKWD-----EVYKIYTNWLA--KVN 722
Qy 978 SH-DFIRNGMANNNAHHQYITQIENNGIIRGQESAGNSVNYKDNPKRSNFSSENHKKN 1036
Db 723 TQIDLIRKMKKE-----ALEN-----QAEATKAIINYQVN--QYTEEEKNNINFN 765
Qy 1037 IQBYNSRDTYRVREIEIKLSK-QNKNNEYSMEYCTYSDERNSSPGPCSRERKLCQCI 1095
Db 766 IDLSSKLNESINKAMININKFLNQCSVYLM-----NSMIPYGVKRLDFDASL 815
Qy 1096 SDYCLKY 1102
Db 816 KDALLKY 822
RESULT 33
US-11-077-550-8
; Sequence 8, Application US/11077550
; Publication No. US2005024435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR FILING DATE: 2002-09-12
; PRIOR FILING DATE: 1999-02-23
; PRIOR FILING DATE: 1997-08-22
; PRIOR FILING DATE: 1996-12-13
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 871
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-11-077-550-8
Query Match 2.0%; Score 129.5; DB 7; Length 871;
Best Local Similarity 19.2%; Pred. No. 0.17;
Matches 163; Conservative 123; Mismatches 322; Indels 239; Gaps 40;
Qy 388 DKNFL---NDLFKNNKNDLDDFFKNE-----KEYDDL-----DC 421

```
Db      83  EKNYLKGVTKLPERIYSTDLGRMLTISVIRGIPFWGGSTIDTELKVIDTNCINVIQPDG 142
QY      422  RYATATIKSFLNGPAKNDVDIASQINVDLRGFCGNYSNNKSNWCT-----GFTTNK 475
Db      143  SYRSEELNLVIIGPSADIIQF-----ECKSFGHEVLNLRNGYSGTQVIRFSPDFTFG 195
QY      476  FPGTCPPRQTLCLGR-----TYLLHRGHEEDYKEHLLGASI----- 513
Db      196  FBSLEVDVTPNPLGAGKFPATPAVTLAHELHAGH-----RLYGAINPNRVFKVNTNA 249
QY      514  -YEAQLLKYYKKE-----KDNALCSIIQNSYADLADI--KGSDI-- 552
Db      250  YEMSGLVSVFBEURLTFGGHDAKFDLSQENFRLLYYNFKDIASTLNKAKSIGVTTAS 309
QY      553  -----KDYGGK---MEENLNKVNKKRNEESLKIFREKWMNDENKENVKMSAVLNKE 605
Db      310  LQYMKNVFKEKYLSSDTSCKFSVDLKFDPKLYKMLTEIYTDNFVKFFKVL-----NRK 364
QY      606  TKDQYD-----KFQKIPQ-----FLRFKEWGGDDFCRKEKIYSPE 642
Db      365  TYLNFDAKVFKNIVPKVNYTIYDGFNLNRLNTLAAFNFGQNTENNMMNFTKLKNFTGLFE 424
QY      643  SFKVEEC-----KKDCDENTCK---NKCSEYKKWIDLKSEYKQVDKYTKDKNKKMY 692
Db      425  FYKLLCVRGIIITSKYSLDKGYNKALNDLCIKVNNW--DLFFSPSE---DNFTDNLNGEE 480
QY      693  DNID-EVKNKEANVYLKESKECKOVNFDKIPNES-PNEYEDMCKCKDEIKYLNKIKYP 750
Db      481  ITSDTNIEAAEENISLDLIQYYLTFNFDNENPENISLENLSSDIIGQLELMP--NIERFP 538
QY      751  KTKHDIYDITFSDTGDGTPISINANINEQQSGKDTSTNGNSETSDSPVSHPEPESDAI 810
Db      539  NGKK--YELDKYT-----MPLYRAQEFHKGSRIALTNS-----VNEALL 577
QY      811  NVEKLSGDESSSTRGILDINDPSTNNVNEVHDASNTQGSVN-TSDITNGHSESSLNR 869
Db      578  NPSRVYTFSSDYVK-----KVNKATEAMFLGWVQLVYDFDTETSE--VST 623
QY      870  TTNAQDIKIRSGNEQSDNOENSHSDNSGSLTIG-----QVPSEDNTQNTY 917
Db      624  TDKIADITIIPIYGPALNIGMLYKDDFVGALIFSGAVILLFPIEIAIP-----VLGTF 679
QY      918  DSQNPHRDTPNALASLPSDDKINEIEGDFDSRDSRSENGRGTNTNTHDVRTNIVSERV 977
Db      680  -----ALVSYIA-NKVLTVQTDIDNALSCKRNEKWD---EVYKYIVTNWLA--KVN 722
QY      978  SH-DFIRNGMANNNAHQYITQIENNGIIRGOESAGNSVNYKDNPKRSNFSSENDHKN 1036
Db      723  TQIDLRKKMKE-----ALEN-----QAEATKAIINYQN--QYTEEEKNNINFN 765
QY      1037  IQBYNGRDTKRVREIIFKLSK-QNKCNNEYSMEYCTYSDBRNSSPGFCSEERKGLCCQI 1095
Db      766  IDLSSKLNESINKAMININKFNQCSVSLM-----NSMIPYGVKRLDFDASL 815
QY      1096  SDYCLKY 1102
Db      816  KDALLKY 822
```

RESULT 34

```
US-11-077-550-26
; Sequence 26, Application US/11077550
; Publication No. US2005024435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Radraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
```

```
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 26
; LENGTH: 871
; TYPE: PRT
; ORGANISM: Clostridium botulinum
; US-11-077-550-26
```

```
Query Match      2.0%; Score 129.5; DB 7; Length 871;
Best Local Similarity 19.2%; Pred. No. 0.17;
Matches 163; Conservative 123; Mismatches 322; Indels 239; Gaps 40;

QY      388  DKNRFU---NDLFKKKNKNDLDDFPKNE-----KEYDLC-----DC 421
Db      83  EKNYLKGVTKLPERIYSTDLGRMLTISVIRGIPFWGGSTIDTELKVIDTNCINVIQPDG 142
QY      422  RYATATIKSFLNGPAKNDVDIASQINVDLRGFCGNYSNNKSNWCT-----GFTTNK 475
Db      143  SYRSEELNLVIIGPSADIIQF-----ECKSFGHEVLNLRNGYSGTQVIRFSPDFTFG 195
QY      476  FPGTCPPRQTLCLGR-----TYLLHRGHEEDYKEHLLGASI----- 513
Db      196  FBSLEVDVTPNPLGAGKFPATPAVTLAHELHAGH-----RLYGAINPNRVFKVNTNA 249
QY      514  -YEAQLLKYYKKE-----KDNALCSIIQNSYADLADI--KGSDI-- 552
Db      250  YEMSGLVSVFBEURLTFGGHDAKFDLSQENFRLLYYNFKDIASTLNKAKSIGVTTAS 309
QY      553  -----KDYGGK---MEENLNKVNKKRNEESLKIFREKWMNDENKENVKMSAVLNKE 605
Db      310  LQYMKNVFKEKYLSSDTSCKFSVDLKFDPKLYKMLTEIYTDNFVKFFKVL-----NRK 364
QY      606  TKDQYD-----KFQKIPQ-----FLRFKEWGGDDFCRKEKIYSPE 642
Db      365  TYLNFDAKVFKNIVPKVNYTIYDGFNLNRLNTLAAFNFGQNTENNMMNFTKLKNFTGLFE 424
QY      643  SFKVEEC-----KKDCDENTCK---NKCSEYKKWIDLKSEYKQVDKYTKDKNKKMY 692
Db      425  FYKLLCVRGIIITSKYSLDKGYNKALNDLCIKVNNW--DLFFSPSE---DNFTDNLNGEE 480
QY      693  DNID-EVKNKEANVYLKESKECKOVNFDKIPNES-PNEYEDMCKCKDEIKYLNKIKYP 750
Db      481  ITSDTNIEAAEENISLDLIQYYLTFNFDNENPENISLENLSSDIIGQLELMP--NIERFP 538
QY      751  KTKHDIYDITFSDTGDGTPISINANINEQQSGKDTSTNGNSETSDSPVSHPEPESDAI 810
Db      539  NGKK--YELDKYT-----MPLYRAQEFHKGSRIALTNS-----VNEALL 577
QY      811  NVEKLSGDESSSTRGILDINDPSTNNVNEVHDASNTQGSVN-TSDITNGHSESSLNR 869
Db      578  NPSRVYTFSSDYVK-----KVNKATEAMFLGWVQLVYDFDTETSE--VST 623
QY      870  TTNAQDIKIRSGNEQSDNOENSHSDNSGSLTIG-----QVPSEDNTQNTY 917
Db      624  TDKIADITIIPIYGPALNIGMLYKDDFVGALIFSGAVILLFPIEIAIP-----VLGTF 679
QY      918  DSQNPHRDTPNALASLPSDDKINEIEGDFDSRDSRSENGRGTNTNTHDVRTNIVSERV 977
```

```

Db 680 -----ALVSYIA-NKVLTVQTDNALSKRNEKWD-----EVYKYIVTNWLA--KVN 722
Qy 978 SH--DFIRNGMANNNAHQYITQIENNGIIRGOESAGNSVNYKDNPKRSNPFSSNDHKKV 1036
Db 723 TQIDLIRKKWKE-----ALEN-----QAEATKAIINYQYN--QYTEERKNINFN 765
Qy 1037 IQEYNSRDTKRVREIILKSK-QNKCNNEYSMEYCTYSDERNSSPGPCSRERKCLCCQI 1095
Db 766 IDLLSSKLNESINKAMININKFLNQC SVLYM-----NSMIPYGVKRLDPDASL 815
Qy 1096 SDYCLKY 1102
Db 816 KDALLKY 822

RESULT 35
US-11-077-550-153
; Sequence 153, Application US/11077550
; Publication No. US2005024435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 153
; LENGTH: 871
; TYPE: PR1
; ORGANISM: Clostridium botulinum
US-11-077-550-153

Query Match 2.0%; Score 129.5; DB 7; Length 871;
Best Local Similarity 19.2%; Pred. No. 0.17;
Matches 163; Conservative 123; Mismatches 322; Indels 239; Gaps 40;

Qy 388 DKNFL---NDLFKNNKNDLDDFFKNE-----KEYDDLCL-----DC 421
Db 83 EKONYLKGVTKLPERIYSTDLGRMLLTSIVRGIPFWGGSTIDTELKVIDTNCINVIQPDG 142
Qy 422 RYTATIISFLNGPAKNDVDIASQINVDLGRGFCNYSKNEKSWNCT-----GTFNKK 475
Db 143 SYRSEELNLVIGGSADIIOF-----ECKSFGHEVLNTRNGYGSTQVIRFSPDFTFG 195
Qy 476 FPGTCEPRPOTLCLGR-----TYLLRHGHEBDYKEHLIGASI----- 513
Db 196 FESLEVDVTPPLGAGKEATDPATVLAHELHAGH-----RLYGIAINRNVFKVNTNA 249
Qy 514 -YRAQLKLYKYE-----KDNALCSIIQNSYADLADII-KGSII----- 552
Db 250 YWMSGLEVSFEELRTGCHDAKFXIDLSQENEFRLYYNFKDIASTLNKAKSIVGTAS 309

```

```

Qy 553 ----KDYGGK----MBENLNKVNKDKRNEESLKIIFREKWMWDENKENVWVMSAVLKNKE 605
Db 310 LQYMKVFKKYLSSDTSKFSVDKLPKDKLYKMLTEIYTEDNFVKFKVL-----NRK 364
Qy 606 TKDYD----KFKQIPQ-----FLWFKWGGDDFCBKRKEKYSFE 642
Db 365 TYLNFQKAVKINIVPKVNYTIYDGFNLRTNLAAFNQONTINNMFYTKJKNFTGLFE 424
Qy 643 SPVEEC-----KKDCDENTCK--NKCSEYKKWIDLKCKSEVKQVDKYTKDKNKKMY 692
Db 425 FYKLLCVRGIIITSKSLDKGYNKALNDLCIKVNW-DLFFSPSE---DNFTDNLKGE 480
Qy 693 DNID-EVYKKBANVYLKSKSKECKVNFDDKIFNES-PNEYEDMCKCKDBIKYLNKIKYP 750
Db 481 ITSDTNIEAAEENISLDLIQQYVLTFFNFONEPENISTENLSSDIIIGQLELMP-NIERFP 538
Qy 751 KTKHDIYDITFSDTGDGTPISINANINEQSGKDTSTNCGNSETSDSPVSHPEPESDAAI 810
Db 539 NGKK--YELDKYT-----MPLYLRAQEFHKGSRIALTNS-----VNSALL 577
Qy 811 NVEKLSGDESSSSTRGILDINDPSVTNNVNEVHDASNTQGSVN-TSDITNGHSESLNR 869
Db 578 NPSRVYTFSSDVVK-----KVNKATEAMFLGWVEQLVVDFTDTESE--VST 623
Qy 870 TTNAAQIKIGRSGNEQSDNQENSSHSDNSGSLTIG-----QVPSEDNTQNTY 917
Db 624 TDKIADITIIPIYGPALNIGNMLYKDDFVGALIFSGAVILLFEIPEIAP----VLGTF 679
Qy 918 DSQNPDRDTPNALASLPDDKINEIGFSSRSRSENGRGTNTNTHDVRNTNIVSERVN 977
Db 680 -----ALVSYIA-NKVLTVQTDNALSKRNEKWD-----EVYKYIVTNWLA--KVN 722
Qy 978 SH--DFIRNGMANNNAHQYITQIENNGIIRGOESAGNSVNYKDNPKRSNPFSSNDHKKV 1036
Db 723 TQIDLIRKKWKE-----ALEN-----QAEATKAIINYQYN--QYTEERKNINFN 765
Qy 1037 IQEYNSRDTKRVREIILKSK-QNKCNNEYSMEYCTYSDERNSSPGPCSRERKCLCCQI 1095
Db 766 IDLLSSKLNESINKAMININKFLNQC SVLYM-----NSMIPYGVKRLDPDASL 815
Qy 1096 SDYCLKY 1102
Db 816 KDALLKY 822

RESULT 36
US-11-077-550-6
; Sequence 6, Application US/11077550
; Publication No. US2005024435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4

```

; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; TYPE: PRT
; LENGTH: 873
; ORGANISM: Clostridium botulinum
US-11-077-550-6

Query Match
Best Local Similarity 19.2%; Score 129.5; DB 7; Length 873;
Matches 163; Conservative 123; Mismatches 322; Indels 239; Gaps 40;

QY 388 DKNFL---NDLFKNNKNDLDDFFKNE-----KEYDDL---DC 421
Db 85 EKONYLKVGTGFERIYSTDLGRMLTISVIRGIPFWGGSTIDTELKVIDTNCINVIQPDG 144

QY 422 RYTATTIKSFLNGPAKNDVDIASQINVDLGRGFCNYKSNKSNWCT-----GTFTNK 475
Db 145 SYRSEELNLVIIGPSADIQF-----ECKSFGHEVLNLRNGYGSQVIRSPDFTFG 197

QY 476 PPTCTPPRRQTLCLGR-----TYLLHGHBEEDYKEHLLGASI----- 513
Db 198 FEESLEVDTNPLLGAGKFATDPAVTLAHLIHAH-----RLYGIAINPNRVFKVNTNA 251

QY 514 -YEAQLLKYYKE-----KDNALCSIIQNSYADLADI--KGSII----- 552
Db 252 YEMSGLEVSFEELRTFGGHDKAFIDSLQENEFRLYYNKFKDIASTLNKAKSIVGTAS 311

QY 553 -----KDYVGGK---MEENLNKVNKKRNEESLKIIFREKWMNDENKENVKMSAVLNKE 605
Db 312 LQYMKNVFKEKYLSEDTSGKFSVDKLPDKLYKMLTEIYEDNFVKFFVL-----NRK 366

QY 606 TCKDYD---KFOKIPQ-----FLRFKMGWDDFCRKRKEKISFE 642
Db 367 TYLNFDAKAVKINIVPKVNTIYDGFNLNTLAANFNGQTEINNWNFTKLNFTGLPE 426

QY 643 SFKVEC-----KKDCDENTCK---NKCSEYKKWIDLKSEYKQVDKYTKDKNKMY 692
Db 427 FYKLLCVRGIIITSKTSLDKGYNKALDLCIKVNNW--DLFFSPSE---DNFTNDLNKGE 482

QY 693 DNID-EVKNKEANVYLKESKECKVDNFDDKIFNES--PNEYEDMCKCKDEIKYLNKYP 750
Db 483 ITSDTNIEAAEENISLDLIQYYLTFNFDNEPENISIEENLSSDIIGGLELMP--NIERFP 540

QY 751 KTKHDIYDIDFTDGDGTPISINANINQOQSGKOTSGNSETSDSPVSHPEPSDAAI 810
Db 541 NGKK--YELDKYT-----MFLYRAQFEGHKSRIALTNS-----VNEALL 579

QY 811 NVEKLSGESSESTRGILDINDPSVTNNVNEVHDASNTQSVSN--TSDITNGHSESLNR 869
Db 580 NPSRVYTFSSDYVK-----KVNKATEAAMFLGWVQLVYDFDTES--VST 625

QY 870 TTNAQDIKIGRSGNEQSDNQSNSHSDNSGSLTIG-----QVPSDNTQNTY 917
Db 626 TDKIADITIIPIYGPAINLGNMLYKDDFVGALIFGCAVILLBFIPIAIP-----VLGTF 681

QY 918 DSQNPHRDTPNALASPSDKINEIEGFSRSDSENGRDTTGNTHDVRRTNIVSERRVN 977
Db 682 -----ALVSYIA-NKVLTVQTDNALSKEKWD---EVYKIYVTNWLA--KVN 724

QY 978 SH-DPIRNGMANNNAHQYITQENNGIIRGOESAGSNVYKDNPKRSNFSSENDHKN 1036
Db 725 TQIDLIRKKWKE-----ALEN-----QAEATKAIINYQN--QYTEEKNNINFN 767

QY 1037 IQSYNGRDTKRVREIIFKLS-QMKCNNEYSMEYCTYSDBRNSPSPGCSREERKLCQOI 1095
Db 768 IDLSSKLNESINKAMININKFLNQCYSYLM-----NSMIPYGVKRLDFDASL 817

QY 1096 SDYCLKY 1102
Db 818 KDALLKY 824

RESULT 37
US-11-077-550-149
; Sequence 149, Application US/11077550
; Publication No. US20050244435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 149
; TYPE: PRT
; LENGTH: 873
; ORGANISM: Clostridium botulinum
US-11-077-550-149

Query Match
Best Local Similarity 19.2%; Score 129.5; DB 7; Length 873;
Matches 163; Conservative 123; Mismatches 322; Indels 239; Gaps 40;

QY 388 DKNFL---NDLFKNNKNDLDDFFKNE-----KEYDDL---DC 421
Db 85 EKONYLKVGTGFERIYSTDLGRMLTISVIRGIPFWGGSTIDTELKVIDTNCINVIQPDG 144

QY 422 RYTATTIKSFLNGPAKNDVDIASQINVDLGRGFCNYKSNKSNWCT-----GTFTNK 475
Db 145 SYRSEELNLVIIGPSADIQF-----ECKSFGHEVLNLRNGYGSQVIRSPDFTFG 197

QY 476 PPTCTPPRRQTLCLGR-----TYLLHGHBEEDYKEHLLGASI----- 513
Db 198 FEESLEVDTNPLLGAGKFATDPAVTLAHLIHAH-----RLYGIAINPNRVFKVNTNA 251

QY 514 -YEAQLLKYYKE-----KDNALCSIIQNSYADLADI--KGSII----- 552
Db 252 YEMSGLEVSFEELRTFGGHDKAFIDSLQENEFRLYYNKFKDIASTLNKAKSIVGTAS 311

QY 553 -----KDYVGGK---MEENLNKVNKKRNEESLKIIFREKWMNDENKENVKMSAVLNKE 605
Db 312 LQYMKNVFKEKYLSEDTSGKFSVDKLPDKLYKMLTEIYEDNFVKFFVL-----NRK 366

QY 606 TCKDYD---KFOKIPQ-----FLRFKMGWDDFCRKRKEKISFE 642
Db 367 TYLNFDAKAVKINIVPKVNTIYDGFNLNTLAANFNGQTEINNWNFTKLNFTGLPE 426

QY 643 SFKVEC-----KKDCDENTCK---NKCSEYKKWIDLKSEYKQVDKYTKDKNKMY 692
Db 427 FYKLLCVRGIIITSKTSLDKGYNKALDLCIKVNNW--DLFFSPSE---DNFTNDLNKGE 482

QY 693 DNID-EVKNKEANVYLKESKECKVDNFDDKIFNES--PNEYEDMCKCKDEIKYLNKYP 750
Db 483 ITSDTNIEAAEENISLDLIQYYLTFNFDNEPENISIEENLSSDIIGGLELMP--NIERFP 540


```
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 159
; LENGTH: 879
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-11-077-550-159

Query Match
Best Local Similarity 2.0%; Score 129.5; DB 7; Length 879;
Matches 163; Conservative 123; Mismatches 322; Indels 239; Gaps 40;

QY 388 DKNFL---NDLKNNKNDLDDFFKNE-----KEYDLC-----DC 421
DB 83 EKONYLKVTKLPERIYSTDLGRMLTSTIVRGIPFWGGSTIDTELKVIDTNCINVIQPDG 142
QY 422 RYATIIKSLFNGPAKNDVDIASQINNDLGRGFCNYKSNKSWNCT-----GFTTNK 475
DB 143 SYRSEELNLVIIGPSADIIQF-----ECKSFGHEVLNLRNGYGSTQVIRFSPDFTFG 195
QY 476 FPGTCPPRRQTLCLGR-----TYLLHRGHEEDYKEHLGLASI----- 513
DB 196 FEESLEVDNPLGLGAGKFPATDPAVLAHELHAGH-----RLYGTAINPNRVFKVNTNA 249
QY 514 -YEAQLLKYYKE-----KDNALCSTIQNSYADLADII-KGSDII----- 552
DB 250 YEMSGLEVSFEELRTFGGHDKAFIDSLQENFRLYYNFKDIASFLNKAISIVGTAS 309
QY 553 ----KDYGKK---MEENLNKVNKKRNEESLKI-FREKWDENKENVKMSAVLKNKE 605
DB 310 LOYMKNVFEKYLSEDTSGKFSVDKLFKDKLYKMLTEIYTDNFVKFFKVL-----NRK 364
QY 606 TCKDYD---KFQKIPO-----FLRWFKWGDDDFCEKREKIYSFE 642
DB 365 TYLNFDAVKPKNIVKPNVTIYDGFNLNLTNLAANFNGQNTNINNFTKLNFTGLFE 424
QY 643 SFKVEC-----KKKDCDENTCK---NKCSEYKKWIDLKKSEYKQVDKYTKDKNKM 692
DB 425 FYKLLCVRGIIITSKTSLDKGYNKALDLCIKVNMW-DLFFSPSE---DNFTNDLNKGE 480
QY 693 DNTD-EVKNKEANYLKEKECKVDNFDKIFNES-PNEYEDMCKCKDIKYLNEIKYP 750
DB 481 ITSDTNTEAABENISLDLIQOYLYLTFFNDEPENISIENTSSDIIGOLELMP--NIBRFP 538
QY 751 KTKHDIDYDIDTDFDGTGDPISINANINEQSGKDTNTGNSSETSDSPVSHPESDAAI 810
DB 539 NGKK--YELDKYT-----MFHYLRAQFEGHKGSRIALTNS-----VNEALL 577
QY 811 NVEKLSGDESSSETRGILDINDPSVTNNVNEVHDASNTQGSVSN--TSDITNGHSESLNR 869
DB 578 NPSRVYTFSSDYVK-----KVNKATEAMFLGWVQVYVDFDTETSE--VST 623
QY 870 TTHAQDTIKGRSGNEQSDNENSHSDNGSLTIG-----QVPSEDNTQNTY 917
DB 624 TDKIADITIIPIYTGPAIINGNMLYKDDFVGALIFSGAVILLEFIPBIAIP-----VLGTF 679
QY 918 DSQNPHRDTPNALASLPDDKINIEGFDSDSRSDENGRGDTTSTNTHDVRRTNIVSBRVN 977
```

```
680 -----ALVSYIA-NKVLTVQTIIDNALSKRNEKWD---EVYKYIVTNMLA--KVN 722
978 SH-DFIRNGMANNNAHQVITQIENNGIIRGOBESAGSNVYKDNPKRSNFSSENHDKN 1036
723 TQIDILIRKWKKE-----ALEN-----QAEATKAIINYQN--QYTEEEKNNINFN 765
1037 IOEYNSRDTKRVREELIKLSK-ONKCNNEYSMEYCTYSDERNSSPQPCSRBEERKKLCCQI 1095
766 IDDLKSLNESINKAMININKFLNQCVSVM-----NSMIPYGVKRLDFDASL 815
1096 SDYCLAY 1102
816 KQALLKY 822

RESULT 40
US-11-077-550-4
; Sequence 4, Application US/11077550
; Publication No. US20050244435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 894
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-11-077-550-4

Query Match
Best Local Similarity 2.0%; Score 129.5; DB 7; Length 894;
Matches 163; Conservative 123; Mismatches 322; Indels 239; Gaps 40;

QY 388 DKNFL---NDLKNNKNDLDDFFKNE-----KEYDLC-----DC 421
DB 106 EKONYLKVTKLPERIYSTDLGRMLTSTIVRGIPFWGGSTIDTELKVIDTNCINVIQPDG 165
QY 422 RYATIIKSLFNGPAKNDVDIASQINNDLGRGFCNYKSNKSWNCT-----GFTTNK 475
DB 166 SYRSEELNLVIIGPSADIIQF-----ECKSFGHEVLNLRNGYGSTQVIRFSPDFTFG 218
QY 476 FPGTCPPRRQTLCLGR-----TYLLHRGHEEDYKEHLGLASI----- 513
DB 219 FEESLEVDNPLGLGAGKFPATDPAVLAHELHAGH-----RLYGTAINPNRVFKVNTNA 272
QY 514 -YEAQLLKYYKE-----KDNALCSTIQNSYADLADII-KGSDII----- 552
DB 273 YEMSGLEVSFEELRTFGGHDKAFIDSLQENFRLYYNFKDIASFLNKAISIVGTAS 332
QY 553 ----KDYGKK---MEENLNKVNKKRNEESLKI-FREKWDENKENVKMSAVLKNKE 605
```


Db 333 LQYMKNVPEKYLLESDTSGKFSVDKLFKXKMLTEIYTEDNFVKFFKVL-----NRK 387
Qy 606 TCKDYD---KFQKIPQ-----FLRWFKEWGDOPCEKKEKIYSFE 642
Db 388 TYLNFDAVKAVKINIVPKNYIYDGFNLRNTNLAANFNGQNTENNWNFTKKNFTGLFE 447
Qy 643 SFKVEEC-----KKKCDENTCK---NKCSEYKKWIDLKSEYKQVDKYTKDKNKKMY 692
Db 448 FYKLLCVRGIIITSKTSKLDKGNKALNDLCIKVNNW-DLPFSPSE---DNFTNDLNKGE 503
Qy 693 DNID-EVKNKEANYLKEKCEKDVNDDKI FNES-PNEYEDMCKCKDEIKYLNEIKYP 750
Db 504 ITSDNTIEAABENISLDLIQQYLYTFNFDNBPENISLENLSSDIIGOLELMP--NIBRFP 561
Qy 751 KTKHDIYDIDTFDGTGTPISINANINEQSGKDTSGTNCSETSDSPVSHPEPESDAAI 810
Db 562 NGKK--YELDKYT-----MFHYLRAQBFHKGKSRALTNS-----VNEALL 600
Qy 811 NVEKLSGDESSETRGILDINDPSTVNNVNEVHDASNTQGSVSN-TSDITNGHSESSINR 869
Db 601 NPSRVYTFSSDYVK-----KVNKATEAAMFLGWVEQLVYDFTDETSE--VST 646
Qy 870 TTNAQDIKIGRSGNEQSDNQENSHSDNSGLTIG-----QVPSEDNTQNTY 917
Db 647 TDRIADITIIIPYIGPALNIGNMLYKODFVGALLIFSGAVILLEFPIETAIIP-----VLGTF 702
Qy 918 DSQNPHRDTPNALASLPDDDKINEIEGDSRDSRSGRGDTTSNTHDVRRTNIVSERRVN 977
Db 703 -----ALVSYIA-NKVLTVQTIDNALSRRNEKWD---EVYKIVTNWLA--KVN 745
Qy 978 SH-DFIRNGMANNAHQYITQIENNGIIRQGESAGNSVNYKDNPKRSNFSSNDHKKN 1036
Db 746 TQIDLIRKKMKE-----ALEN-----QAEATKALINYQYN--QYTEEEKNNINFN 788
Qy 1037 IQEYNSRDTKRVREIILKSK-QNKCNNEYSMEYCTYSDERNSSPGPCSRERKKLCCQI 1095
Db 789 IDDLSSKLNESINKAMININKFLNQCSVSYLM-----NSMIPYGVKRLDFDASL 838
Qy 1096 SDYCLKY 1102
Db 839 KDALLKY 845

Search completed: November 21, 2005, 20:52:45
Job time : 30 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2005, 20:17:17 ; Search time 193 Seconds
(without alignments)
2754.655 Million cell updates/sec

Title: US-10-677-980-2

Perfect score: 6481

Sequence: 1 MKGYFNLYFLPLFLYLVNI.....VOETNISDSYNNYNNKMY 1210

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_21:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*
- 9: Geneseq2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6481	100.0	1210	6	ABG73547 P. falcip
2	6081	93.8	1143	5	Aau76759 Plasmodiu
3	1557	24.0	1435	2	AAR70232 P. falcip
4	1557	24.0	1435	2	Aaw22477 Silaia ac
5	1557	24.0	1435	3	Aay77900 P. falcip
6	1543.5	23.8	1604	2	Aar70105 TNF-R-EBA
7	1542.5	23.8	1786	2	AAR41043 CD4-EBA17
8	1537.5	23.7	1421	5	Aau76764 Plasmodiu
9	1200.5	18.5	616	6	ABR82499 Synthetic
10	1199.5	18.5	616	6	ABR82498 Synthetic
11	1197.5	18.5	616	5	AAM50533 Unidentif
12	1112	17.2	1086	5	Aau76760 Plasmodiu
13	1088	16.8	1568	6	ABP70152 Amino aci
14	993.5	15.3	1501	5	Aau76762 Plasmodiu
15	768	11.9	1115	2	AAR13457 Duffy rec
16	761.5	11.7	749	2	AAR70233 P. falcip
17	761.5	11.7	749	2	Aaw22479 Plasmodiu
18	761.5	11.7	749	3	Aay77901 P. falcip
19	754	11.6	1115	2	Aaw22478 Duffy ant
20	754	11.6	1115	3	Aay77899 P. vivax
21	750.5	11.6	972	5	Aau76761 Plasmodiu
22	743.5	11.5	1245	5	AAR70106 TNF-R-Pl.
23	704.5	10.9	1061	2	AAR70231 P. vivax
24	698	10.8	1028	2	AAR41044 Plasmodiu

25	684	10.6	350	5	ABB07656	Abb07656 P. falcip
26	655	10.1	778	2	AAR13456	Aar13456 Duffy rec
27	651	10.0	2459	8	ADO69869	Ado69869 Plasmodiu
28	637.5	9.8	3056	8	ADR43499	Adr43499 CSA-Bindi
29	601.5	9.3	3542	4	AAB62142	Aab62142 P. falcip
30	592	9.1	3147	9	AEb22180	Aeb22180 Codon opt
31	592	9.1	3553	9	AEb22174	Aeb22174 Plasmodiu
32	554.5	8.6	2994	9	AEb22176	Aeb22176 Plasmodiu
33	546.5	8.4	302	5	ABB07655	Abb07655 P. falcip
34	546.5	8.4	2913	2	AAW00384	Aaw00384 Plasmodiu
35	543.5	8.4	3346	9	AEb22178	Aeb22178 Plasmodiu
36	498.5	7.7	3060	2	AAW22475	Aaw22475 Plasmodiu
37	498.5	7.7	3060	3	AAW22475	Aaw22475 Plasmodiu
38	495.5	7.6	2703	2	AAW77905	Aay77905 Plasmodiu
39	495.5	7.6	2710	2	AAW22482	Aaw22482 Plasmodiu
40	495.5	7.6	2710	3	AAW77904	Aay77904 P. falcip
41	452.5	7.6	697	9	ADZ72259	Adz72259 Plasmodiu
42	432	6.7	1507	3	AAW24128	Aaw24128 Plasmodiu
43	414	6.4	762	9	ADZ72267	Adz72267 Plasmodiu
44	410.5	6.3	2485	3	AAW18172	Aaw18172 Plasmodiu
45	409.5	6.3	2647	8	ADP25446	Adp25446 Plasmodiu

ALIGNMENTS

RESULT 1
ABG73547

ID ABG73547 standard; protein; 1210 AA.

XX AC ABG73547;

XX DT 24-FEB-2003 (first entry)

XX DE P. falciparum BAEBL protein.

XX KW BAEBL; erythrocyte binding protein; protozoacide; immunostimulant;

XX KW malaria; parasite; vaccine; chromosome 13.

XX OS Plasmodium falciparum.

XX FH Key Location/Qualifiers

XX FT Misc-difference 185

XX FT /note= "A polymorphism resulting in variation from Val to Ile at this position is specifically claimed in Claim 8"

XX FT Misc-difference 239

XX FT /note= "A polymorphism resulting in variation from Ser to Asn at this position is specifically claimed in Claim 8"

XX FT Misc-difference 261

XX FT /note= "A polymorphism resulting in variation from Thr to Arg at this position is specifically claimed in Claim 8"

XX FT Misc-difference 285

XX FT /note= "A polymorphism resulting in variation from Lys to Glu at this position is specifically claimed in Claim 8"

XX PN WO200278603-A2.

XX PD 10-OCT-2002.

XX PP 29-MAR-2002; 2002WO-US010071.

XX PR 02-APR-2001; 2001US-0281130P.

XX FA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Mayer G, Miller LH;

XX XX WPI; 2003-092869/08.

XX DR N-PSDB; ABG73547.

XX XX New vaccine against malaria Plasmodium falciparum parasite comprising Erythrocyte Binding Protein polypeptide.

CC domain polypeptide. The invention is further useful for research
CC applications, vaccine candidates, blocking peptides, diagnostics and
CC prognostics, where diseases of Plasmodium falciparum include malaria. The
CC polypeptide of the invention and antibodies are useful for diagnosis of
CC malaria, and for detection of P. falciparum in culture media and in
CC biological samples such as biological tissues and fluids. The present
CC amino acid sequence represents the Plasmodium falciparum EBP2 protein
CC which is one of several (AAU76759-AAU75762) EBA-175 paralogues
xx
SQ Sequence 1143 AA;

Query Match				93.8%	Score 6081;	DB 5;	Length 1143;						
Best Local Similarity				99.6%;	Pred. No. 0;								
Matches 1130; Conservative				3;	Mismatches 1;	Indels 0;	Gaps 0;						
Qy	1	MKG	YFNIFLPIPLIFLVNIRINESII	IGRTL	YNRODESSDI	SRVNSPELANNHKTNI	YDS 60						
Db	7	MKG	YFNIFLPIPLIFLVNIRINESII	IGRTL	YNRODESSDI	SRVNSPELANNHKTNI	YDS 66						
Qy	61	DYED	VNNKLNSFVENKSVKGRSL	FINNKT	KSYDI	IIPPSYSYRNDKFNLSSE	EDNSG 120						
Db	67	DYED	VNNKLNSFVENKSVKGRSL	FINNKT	KSYDI	IIPPSYSYRNDKFNLSSE	EDNSG 126						
Qy	121	NTNSNN	PANTSEISIGKDNKQYTFI	QKRTHL	FACGIRKSI	KIKWICRENSK	ITVCPDRK 180						
Db	127	NTNSNN	PANTSEISIGKDNKQYTFI	QKRTHL	FACGIRKSI	KIKWICRENSK	ITVCPDRK 186						
Qy	181	IOLCV	ANFLNSRLTMEKFEI	PLISVNT	EAKLLYNNEG	KDPSIFCNELRNS	SPDPRSS 240						
Db	187	IOLCV	ANFLNSRLTMEKFEI	PLISVNT	EAKLLYNNEG	KDPSIFCNELRNS	SPDPRSS 246						
Qy	241	FIGDDM	FGGNTORVKGYINTK	SDYYKEKN	VEKLN	IKKEWEKKNANLWNH	IVNHNK 300						
Db	247	FIGDDM	FGGNTORVKGYINTK	SDYYKEKN	VEKLN	IKKEWEKKNANLWNH	IVNHNK 306						
Qy	301	NISKE	CAIIIPAE	BPQINLWIK	EWNEFLME	KRLFLNIKDKCVENK	YKYEACFGGCR	LPCS 360					
Db	307	NISKE	CAIIIPAE	BPQINLWIK	EWNEFLME	KRLFLNIKDKCVENK	YKYEACFGGCR	LPCS 366					
Qy	361	SYTS	FMKSKTQME	VLNLYKCKNS	GVDXNN	FLNDLFPKNNKND	LDDFFFKNEKEY	DDLCD 420					
Db	367	SYTS	FMKSKTQME	VLNLYKCKNS	GVDXNN	FLNDLFPKNNKND	LDDFFFKNEKEY	DDLCD 426					
Qy	421	CRYT	ATTIIKSFL	NGPAKNDV	DIASQIN	VNDLRGFGCNYKSN	NEKSWNCTGTFTN	KFPGTC 480					
Db	427	CRYT	ATTIIKSFL	NGPAKNDV	DIASQIN	VNDLRGFGCNYKSN	NEKSWNCTGTFTN	KFPGTC 486					
Qy	481	EPPE	RQTLCLGR	TLYLLHRG	HEEDYK	HLGLASIYEAQLLYKYK	EKEDENALCS	TIQNSYA 540					
Db	487	EPPE	RQTLCLGR	TLYLLHRG	HEEDYK	HLGLASIYEAQLLYKYK	EKEDENALCS	TIQNSYA 546					
Qy	541	DLADI	IKGSDII	IKDYYG	KQWEENLKN	YKNDKCRNE	ESLKI	PREKWM	DENKENVKN	MSAV 600			
Db	547	DLADI	IKGSDI	IKDYYG	KQWEENLKN	YKNDKCRNE	ESLKI	PREKWM	DENKENVKN	MSAV 606			
Qy	601	LKN	KETCKOYDK	FQKI	POFLR	WPKEWGD	DFCEK	REKEKI	YSPES	PKVECKK	CCDENTCKN 660		
Db	607	LKN	KETCKOYDK	FQKI	POFLR	WPKEWGD	DFCEK	REKEKI	YSPES	PKVECKK	CCDENTCKN 666		
Qy	661	KCSEY	KKWIDLK	KSEY	EKQVD	KYTKDK	KNKQWYDNI	DEVQ	NKEANV	YLKEK	SKCEKQV	NFD 720	
Db	667	KCSEY	KKWIDLK	KSEY	EKQVD	KYTKDK	KNKQWYDNI	DEVQ	NKEANV	YLKEK	SKCEKQV	NFD 726	
Qy	721	DKI	FNES	PNEY	EDMCK	KDEIK	YLNEIKY	PKTKH	DIYDID	FTSD	FGDGP	PISINANINE 780	
Db	727	DKI	FNES	PNEY	EDMCK	KDEIK	YLNEIKY	PKTKH	DIYDID	FTSD	FGDGP	PISINANINE 786	
Qy	781	QOS	GKOT	NTNG	SETSD	SPVSH	PESDAA	INVEK	LSGDES	SETR	GIL	DINDPSV	TNNVN 840
Db	787	QOS	GKOT	NTNG	SETSD	SPVSH	PESDAA	INVEK	LSGDES	SETR	GIL	DINDPSV	TNNVN 846
Qy	841	EVHD	ASNTQ	GSV	NTSD	ITNGH	SESSLN	RTTNA	QDIKIG	RSNGE	QSDNS	QNSHSS	DNSG 900

Db	847	EVHDASNTQGSVNTSDITNGHSESSLNRTTNAQDIKIGRSGNEQSDNQENSHSHSDNSG	906
Qy	901	SLATIGQVPSEDNTQNTYDSQNPHRDTPNALASLPDDKLINEIEGFSSRDSSENGRGDTTS	960
Db	907	SLATIGQVPSEDNTQNTYDSQNPHRDTPNALASLPDDKLINEIEGFSSRDSSENGRGDTTS	966
Qy	961	NTHDVTRTNVTSERRVNSHDFIRNGMANNNAHHQYITQIENNGIIRGQESAGNSVNYKD	1020
Db	967	NTHDVTRTNVTSERRVNSHDFIRNGMANNNAHHQYITQIENNGIIRGQESAGNSVNYKD	1026
Qy	1021	NPKRSPFSSNDHKNIQEVNSRDTKRVREELIKLQKQNCNNEYSMEYCTYSDERNSSP	1080
Db	1027	NPKRSPFSSNDHKNIQEVNSRDTKRVREELIKLQKQNCNNEYSMEYCTYSDERNSSP	1086
Qy	1081	GPCSRERKKLCCQISDYCLKYFNFYISIEYNCIKSEIKSPYKCFKSEGOSSI	1134
Db	1087	GPCSRERKKLCCQISDYCLKYFNFYISIEYNCIKSEIKSPYKCFKSEGOSSM	1140
RESULT 3			
AA70232			
ID	AA70232	standard; protein; 1435 AA.	
XX	AC	AA70232;	
XX	DT	25-MAR-2003 (revised)	
DT	22-SEP-1995	(first entry)	
XX	DE	P. falciparum SABP.	
XX	KW	SABP; sialic acid binding protein; binding domain; merozoite; malaria; therapy; vaccine.	
XX	OS	Plasmodium falciparum.	
XX	FH	Key Location/Qualifiers	
FT	Domain	1..616	
FT		/label= Binding_domain	
XX	PN	WO9507353-A2...	
XX	FD	16-MAR-1995.	
XX	PF	07-SEP-1994; 94WO-US010230.	
XX	PR	10-SEP-1993; 93US-00119677.	
XX	PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.	
XX	PI	Sim Kl, Chitnis C, Miller LH, Peterson DS, Su X, Wellemes TE;	
XX	DR	WPI; 1995-123427/16.	
XX	DR	N-PSDB; AAQ83525.	
XX	PT	New erythrocyte binding domain polypeptide(s) - isolated from Plasmodium	
XX	PS	binding proteins, used in diagnosis, treatment and prevention of malaria.	
XX	PS	Disclosure; Page 39-41; 81pp; English.	
XX	CC	Sequences from the SABP gene (given in AAQ83525) were PCR amplified,	
XX	CC	expressed on the surface of COS cells and tested for erythrocyte binding	
XX	CC	to identify the binding domain polypeptide. A pref. SABP binding domain	
XX	CC	comprises residues 1 to about 616 of the SABP protein (AA70232).	
XX	CC	Recombinant binding domain was expressed in E. coli, yeast, mammalian,	
XX	CC	insect, and in vaccinia virus and adenovirus-infected cells. It provides	
XX	CC	protection against P. falciparum. (Updated on 25-MAR-2003 to correct PN	
XX	CC	field.)	
XX	SQ	Sequence 1435 AA;	
Query Match		24.0%; Score 1557; DB 2; Length 1435;	
Best Local Similarity		28.5%; Pred. No. 2e-88;	
Matches	410; Conservative	214; Mismatches 467; Indels 348; Gaps 42;	

CC	Binding Protein (SABP), which are soluble proteins that appear in the culture supernatant after erythrocytes infected with malaria release merozoites. Immunochemical studies indicate that DABP and SABP are the respective ligands for Plasmodium vivax and Plasmodium falciparum Duffy and sialic acid receptors on erythrocytes. The ebl-1 polypeptides may be used to vaccinate against malaria, especially caused by P. falciparum.	
CC	Immunization with the polypeptide provides effective protection against malaria. The present sequence represents the SABP binding domain polypeptide	
XX		
SQ	Sequence 1435 AA;	
	Query Match 24.0%; Score 1557; DB 3; Length 1435; Best Local Similarity 28.5%; Pred. No. 2e-88; Matches 410; Conservative 214; Mismatches 467; Indels 348; Gaps 42;	
QY	1 MKGYFIYFIPLIFLYNVRINESIIGRTLYNRQDESSDISRVSPELNNHKTNIYDS 60	
DB	1 MKCNISIIYFPASFVLYFAKARNEYDI-----KENEFLDVYKEFNELDKKYGVNQKT 55	
QY	61 DYEDVNNKLINSFVENKSVKKKSLSPINN-----KTKSY---DIIPPSYSYRNDKPNSL 112	
DB	56 D-----KKIFTFTENK-----LDILNNSKFNKRWKSYGTPDNIDKNMSLIN----- 96	
QY	113 SENEDSGNTNSNPANTSEISIGKONKQ-----YTFIQKETH----- 150	
DB	97 KRNNEEMFNNSYOSFUSTS--SLIKONKYYPINAVRVSRILSFLDSRINNGRNTSSNNEV 154	
QY	151 LFACGIRKRSIKWICRENSKIT--VCVPDRKIQLCVANFLNSLRTWEKPEKFIPLISVNT 209	
DB	155 LSNCREKRKGMKWDCKKNDPSYVCIPDRRIQLCVNLISIIKTYTKEWKDHFIEASKK 214	
QY	210 EAKLLYNNKSGKDPISPCNBLRNSFSDFRSFIDGDMDFCGNTDRVKGYINTKPSDYIKE 269	
DB	215 ESQLLAKNDKNYKSKPCNDLKNKSLFDYCHLAWGNDMDFGGYSTKAENKIQEVFKGAHGE 274	
QY	270 KNVEKLNNIKEWKEKNKANLWHMIVNHHKGNISKECAIIPAREPQINLWIKWENENFLM 329	
DB	275 ISEHKIKNFRKWNNEFKLEWAMLSEHKNNIN--NCKNIPQBELQITQWIKWHGEFL 333	
QY	330 EKRLFLNIDKVENKKEACFGGCRLLPCSSYTSFMKSKTQMEVLITNLYKKKNSGVDK 389	
DB	334 ERDNRSKPKSKKNNTLYEACEKECIDPCWKYRDMIRSKFEWHILSKEYETQK--VPK 391	
QY	390 NNFENLDFK--KNNQND--LDDFPN--EKBYDDLDLCDCRYTATIIKSFINGPAKNVDIASQ 445	
DB	392 ENAENYLIKISENKDAKVSLLANNCDAYSKEYCDCKHITTLVKSVLNGNDNTIKEKREH 451	
QY	446 INVNDLRGFCGNYS--NNEKSNWCTGTFTNKPFGTCPEPRROTCLGRTYLLHRGHEED 503	
DB	452 IDLDGFSKFGCDKNKSVDTNTKWECKRNPYILSTKDVCCVPPRQBELCLGNIDRIYKNLLM 511	
QY	504 YKEHLLGASTIYEAQLLYKYEKEDENALCSIIIONSVADLADIIGSDIHKDYVKGKMEEN 563	
DB	512 IKEHILAIAIYESRILKRYKNKDDRECKVIKNTFADIRDIIGCTDYNDLSNRKLVGK 571	
QY	564 LNKVNDKKNRNESSLKIFREKWDENKENYKVMASAVLNKETCKDYDKFKQKIPQFLRWF 623	
DB	572 INTNSKYVHRNKKNDKLFREDWVKVIKKDVWNVISWVFKDVTCKE--DDIENIPQFPRWF 630	
QY	624 KENGDDFCERKEKIKIYSFSPKVECKKCDKDCENTCKNKCEYKKWIDLKSEYEKQV--- 680	
DB	631 SEWGDDYCDQKTKMI-----ETLKVECKEKPCEDDNCKSKCNYSKEWISKCKEYNNQAKQY 687	
QY	681 DKYTKDNKKKQYDNIDVQNKKEANVYLKESKECKDVNFDDKIFNESPNEYEDMCKKCD 740	
DB	688 QEYQKGNYYKY---SEFKSIKEVLYLKYSEKSNLNFEDFEKELHSDYKKNKCTWCPE 744	
QY	741 IKYL-----NE----- 746	
DB	745 VKDVPISIIIRNEQTSQEAPEENTEIAHRTETPSISEGPKGNEQKERDDDSLKISVSP 804	
QY	747 -----IKYPKT----- 752	

DB	805 ENSRPTDAKQTSNLLKLKGDVDIMPKAVIGSSPNNDINVTQGGDNISGVNSKPLSDDV 864	
QY	753 ---KHDIIYID-----TFSDTTFGDGTPISI----- 774	
DB	865 RPKKLEQNDSESEETVNVNHIKSPSINNNGDDSGSGSATVSESSSNTGLSIDDDRNG 924	
QY	775 -----NANINEQOSKQTSNTGNSETSDSP---VSHPEPSDAAINV 812	
DB	925 DTFVTRQDTANTEDVIRKENADKDEKGADEERHSTSELSPEEKMLTLDNEGNSLNH 984	
QY	813 EKL---SGDESSETRGILDIN-----DPSVTNNVNEVH-----DAS 846	
DB	985 EEVKEHTSNDNVQSGGIVNMNVEKELKDTLENPSSLDDEGKAHELESEPNLSSDDMS 1044	
QY	847 NTQGSVNTSDITN---GHSESLAR-----TTNAQDIKIGRSGNEOSDN---QENS 892	
DB	1045 NTPGPLDNTSEETTERISNNEYKVNEREDERTLTKEYEDIVLKSHMNRSDDCGELYDENS 1104	
QY	893 SHSSDNGSLTI--GVVPSDNTQNTY-----DSQNPHRDTPNALA-----SLP- 934	
DB	1105 DLSTVNDSESEDAEAKMGNDTSEMHSNSSHIESDOQKNDMKTVDGLGTHVQNEISVPV 1164	
QY	935 ---SDDKINE-----IEGFSRSRDSSENGRGDTTSNTHDVR-----R 967	
DB	1165 TGEIDKLRSESKESKIHKAEBERLSHTDTHKINPEDRNSNTLHLKDIRNEENRHLTNQN 1224	
QY	968 TIVTSERRVNSHDFIRNGMANNNAHQYITQ-----IENNGIIRGOESAGNSVNYKDN- 1021	
DB	1225 INISQERDIQKHGF---HTMNNLHGDGVSERSQINSHSHGNFQDRGSGENVLWRSNN 1280	
QY	1022 -----PKSNTSSSENDHKNIQEVNSRDTKRVREIIKLKSKONKCNNEYSMEYCTYSDE 1075	
DB	1281 NNFNNIPSRYNL---YDKKLDLDLYENRNDSTTKELIKKLAELKNCENEISVKYCDHMIH 1337	
QY	1076 RNSSPGPCSPREERKKLCCOISDYCLKYNFYSTEVYNCIKSEIKSPYKCFKSEGOSSI 1134	
DB	1338 BEIPLKTCCTKEKTRNLCCAASDYCMYSFTYDSEYYNCTKREPDDPSYTCFRKEAFSSM 1396	
	RESULT 6	
	AAR70105	
ID	AAR70105 standard; protein; 1604 AA.	
XX	AAR70105;	
AC		
DT	16-OCT-2003 (revised)	
DT	25-MAR-2003 (revised)	
DT	02-NOV-1995 (first entry)	
XX		
DE	TNF-R-EBA 175 fusion protein.	
XX		
KW	Hybrid peptide; malaria parasite; plasmodium falciparum; fusion protein; red blood cell; cytokine receptor; glycoporin binding peptide 130;	
KW	GBP 130; GBPH; glycoporin binding peptide homologue; glycoporin A.	
XX		
OS	Homo sapiens.	
OS	Plasmodium falciparum.	
OS	Chimeric.	
XX		
PN	WO9506737-A1.	
XX		
PD	09-MAR-1995.	
XX		
PF	01-SEP-1994; 94WO-GB001900.	
PR	03-SEP-1993; 93GB-00018350.	
PR	23-AUG-1994; 94GB-00017021.	
XX	(PREN/) PRENDERGAST K F.	
PA	Prendergast KF;	
XX		
PI		
XX		


```

RESULT 8
AAU76764
ID AAU76764 standard; protein; 1421 AA.
XX
AC AAU76764;
XX
DT 21-MAY-2002 (first entry)
XX
DS Plasmodium falciparum erythrocyte binding protein EBA-175 (EBP1).
XX
KW Erythrocyte binding protein; EBP; malaria parasite; protozoacide;
KW vaccine; immune response inducer; Plasmodium falciparum merozoite; SAPP;
KW salicylic acid binding protein; EBA-175.
XX
OS Plasmodium falciparum.
XX
FH Key Location/Qualifiers
FT Region 20..157
FT /label= Region_I
FT Region 145..760
FT /label= Region_II
FT /note= "Region II is subdivided into regions F1 and F2"
FT 743..1322
FT /label= Region_III_V
FT Region 1304..1394
FT /label= Region_VI
XX
PN WO200211756-A2.
XX
XX 14-FEB-2002.
XX
PF 07-AUG-2001; 2001WO-US024725.
XX
XX 07-AUG-2000; 2000US-0223525P.
XX
XX (ENTR-) ENTREMED INC.
XX
XX Narum DL, Sim KL;
XX
XX WPI; 2002-227117/28.
XX
XX Paralog of erythrocyte binding protein-175 polypeptide sequence useful
XX for inducing immune response to Plasmodium falciparum merozoites, that
XX causes malaria, in a patient, and for diagnostic and prognostic purposes.
XX
XX Disclosure; Fig 2; 52pp; English.
XX
XX The present invention relates to a new polypeptide comprising a paralogue
XX of erythrocyte binding protein-175 (EBA-175) polypeptide sequence. The
XX invention is useful for inducing an immune response to Plasmodium
XX falciparum merozoites in a patient. The method of the invention comprises
XX administration of isolated SAPP (salicylic acid binding protein) binding
XX domain polypeptide. The invention is further useful for research
XX applications, vaccine candidates, blocking peptides, diagnostics and
XX prognostics, where diseases of Plasmodium falciparum include malaria. The
XX polypeptide of the invention and antibodies are useful for diagnosis of
XX malaria, and for detection of P. falciparum in culture media and in
XX biological samples such as biological tissues and fluids. The present
XX amino acid sequence represents the Plasmodium falciparum EBA-175 protein
XX of the invention. This sequence is also known as erythrocyte binding
XX protein EBP1
XX
SQ Sequence 1421 AA;
XX
Query Match 23.7%; Score 1537.5; DB 5; Length 1421;
Best Local Similarity 28.1%; Pred. No. 3.3e-87;
Matches 411; Conservative 211; Mismatches 467; Indels 375; Gaps 44;
XX
XX 1 MKGVNLYELPLFLYNVIRINESIGRTLYNQDESSDISRVNSPELNNHKTNIYDS 60
XX ||| :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| :|
XX 1 MKCNISYFFASFFVLYFAKARNEYDI-----KENEKFLDYVYKEKFNELDKKYGNYQKT 55
XX

```

Db 1105 TKEVEDIVLASHNRSDGDELDYDENSDLSTVNDSEDAEAKMGKNDTSEMSHNSOHE 1164
 Qy 918 -DSQPHRDTPNALA-----SLP-----SDDKINE-----IEGFDSSRDS 951
 Db 1165 SDQKNDMKTGDLGTTTHVQNEISVPTGTBIDEKLRSEKSKTHKABEERLSHTDIHKIN 1224
 Qy 952 ENRGDGTTSNTHDVR-----RTNIVSERRVNSHDFIRNGMANNNAHQYITQ-- 998
 Db 1225 PEDRNSNTHLKHDKIQRNENRHLTNQINISQBRDLQKHGF-----HTMNNLHGDGVRS 1280
 Qy 999 ---IENNGIIRGOESAGNSVNYKDN-----PKRSNFSSSENDHKKNIQEYNSRDTKRV 1048
 Db 1281 QINHSHHGNGQDRGSGNSVNLWRSNNNNFNIPSRVNL---YDKKLDLDLYENRDSIT 1337
 Qy 1049 REBIIKLSKQNKNNVSMYCYTYSDBRNSSPQPCSRERKLCQOISDYCLKYFNFYSI 1108
 Db 1338 KELIKKLAELINKCENEISVKYCDHMIHEEIPLTKCTKEKTRNLCVADYCMSTFTYDSE 1397
 Qy 1109 EYNYCKISEIKSPYCKFKSEGS 1132
 Db 1398 EYNYCKREFDPSYTCFRKEAFS 1421

RESULT 9

ABR82498

ID ABR82499 standard; protein; 616 AA.

XX AC ABR82499;

XX DT 20-NOV-2003 (first entry)

XX DE Synthetic erythrocyte-binding protein EBA-175 RII sequence.

XX KW BFP; erythrocyte-binding protein; malaria; pathogen; pProtozoaside;

XX KW immunostimulant; vaccine; EBA-75; EBA-175 RII.

XX OS Synthetic.

XX OS Plasmodium falciparum.

XX XX WO2003062374-A2.

XX XX 31-JUL-2003.

XX XX 12-NOV-2002; 2002WO-US036368.

XX XX 09-NOV-2001; 2001US-0345051P.

XX XX (ENTR-) ENTREMED INC.

XX PI Narum DL, Liang H, Fuhrmann S, Sim K;

XX XX WPI; 2003-598746/56.

XX XX DR N-PSDB; ACF35993.

XX XX PT New synthetic genes encoding erythrocyte-binding protein of Plasmodium

XX XX falciparum, useful as immunization agents for inducing anti-malarial

XX XX response and/or treating malaria, or for raising antibodies for passive

XX XX immunization.

XX XX Claim 5; Fig 9; 73pp; English.

XX XX The invention relates to new synthetic nucleotide sequences (I) encoding

XX CC erythrocyte-binding protein (EBP) of a malaria pathogen, where codon

XX CC usage of the synthetic nucleotide sequence is altered compared to a

XX CC naturally occurring sequence of the EBP to approximate codon usage of a

XX CC host of the malaria pathogen. (I) and the purified erythrocyte binding

XX CC protein are useful for the production of a medicament for induction of an

XX CC anti-malarial immune response in a human or animal. The methods are

XX CC useful for inducing an anti-malarial immune response in a host of the

XX CC malarial pathogen, optimizing expression in P. pastoris of a recombinant

XX CC Plasmodium falciparum region II EBA-75 erythrocyte-binding protein, and

XX CC obtaining the pure protein, and obtaining a DNA vaccine for inducing an

CC anti-malarial immune response. (I) is useful in the production of anti-
 CC malaria DNA vaccines with improved immunogenicity, in expression vectors
 CC for the production of malaria EBPs, or as an immunization agent for
 CC inducing anti-malarial response and/or treatment of malaria. The EBPs are
 CC useful for raising antibodies or serum for passive immunization and
 CC treatment of malarial disease. The present sequence represents a EBA-175
 CC RII synthetic protein without N-glycosylation sites
 CC SQ Sequence 616 AA;

Query Match 18.5%; Score 1200.5; DB 6; Length 616;
 Best Local Similarity 38.4%; Pred. No. 1.5e-66;
 Matches 246; Conservative 111; Mismatches 236; Indels 47; Gaps 13;

Qy 154 CGIKRSIKIKWCRENSEKIT-VCPDRKIQLCVANFLNSRLETMEKFKETPLISVNTPEAK 212
 Db 14 CREKRGKMWCKCKKNDKSNVYVCPDRKIQLCVANFLNSRLETMEKFKETPLISVNTPEAK 73
 Qy 213 LLYNKGEGKDPISFCNELRNSDFRSSFIDGMDFGGNTDRVKGYNITKFSYDYEKNV 272
 Db 74 LLLKKNNDKYNKFCNDLKNSEFLDYGHLAGNDMDFGGYSTKAENKIQEVFKGAHGEISE 133
 Qy 273 EKLNNIKKEWKEKNKANLWNHMIVNHKNISKECAIIPAEPOINLWIKWENNENFLMEKK 332
 Db 134 HKIKFRKKWNEFREKLMSEHKNIN-NCKNIPQELQITQWIKWHGEFLIERD 192
 Qy 333 RLFLNITKCKVENKYEACFGGRLPCSSYTPSMKSKTQMEVLTNLYKCKNSGVDKNNF 392
 Db 193 NRAKLPKSKCNALYEAKECEIDPCMKYRDWIRSKPEWHTLSKEYETQK--VPKENA 250
 Qy 393 LNDLFK-KNNKD--LDDFFKN-EKEYDDLCDRYTATIKSFLNGPAKNDVDIASQINV 448
 Db 251 ENYLKISENKONDAKVSLLNNDCAEYSKYCDCKHTTLVKSVLNGNDNTIKEREHIDL 310
 Qy 449 NDLRGFGCNVKS--NNEKSWNCTGTFTNKPFGTCEPPRQTLCLGRYLLHRGHEEDYE 506
 Db 311 DDFSKEGCDKNSVDNTWKVECKPYKLTQVCPVPRQELCLGNIDRIYDKNLLMIKE 370
 Qy 507 HLGASIVBAQLKYYKKEKDENCALSIIONSADLADIIKGSIIIDKYYGKKMEENLNK 566
 Db 371 HILAIAYESRLKRYKNKDDKEVCKIIQKTADIRDIIGGTDYWNLSNRKLVGKINT 430
 Qy 567 VNKDKKEEESLKIPEKWDENKENVKMSAVLNKKECTCKDYDFQKIPQLRWFKEW 626
 Db 431 NSNYHRNKQNDKLFDEWVKVIKDVWNVISWVFKDKTVCKE-DDIENIPQFRWFSEW 489
 Qy 627 GDDFCERKREKIYSPBSFKVECKKDCDENTCKNCKSEYKKWIDLKSEYEKQV---DKY 683
 Db 490 GDDYCDQKTMI---ETLKECKEKECEDDNCCKRCKNSYKWKSEYKKEEYNKQAKQYQY 546
 Qy 684 TKDKNKKMYDNIDEVKNKEANVYLKESKCEKDVNFDDKIFNESPNEYEDMCKKCEIKY 743
 Db 547 QKGNMYKMY---SEPKSIKPEVYLKYSKCSNLNFEDEPKELSHSDYKKNCTMCPVK- 602
 Qy 744 LNEIKPKTKHDIYDIDTFSDTFGDTGPISINANINEQOS 783
 Db 603 -----DVPISITIRN-NEQTS 616

RESULT 10

ABR82498

ID ABR82498 standard; protein; 616 AA.

XX AC ABR82498;

XX DT 20-NOV-2003 (first entry)

XX DE Synthetic erythrocyte-binding protein EBA-175 RII.

XX KW BFP; erythrocyte-binding protein; malaria; pathogen; pProtozoaside;

XX KW immunostimulant; vaccine; EBA-75; EBA-175 RII.

XX OS Synthetic.

OS Plasmodium falciparum.
 XX WO2003062374-A2.
 XX PD 31-JUL-2003.
 XX PF 12-NOV-2002; 2002WO-US036368.
 XX PR 09-NOV-2001; 2001US-0345051P.
 XX PA (ENTR-) ENTREMED INC.
 XX PI Narum DL, Liang H, Fuhrmann S, Sim K;
 XX DR WPI; 2003-598746/56.
 XX DR N-PSDB; ACF35992.
 XX
 XX New synthetic genes encoding erythrocyte-binding protein of Plasmodium
 PT falciparum, useful as immunization agents for inducing anti-malarial
 PT response and/or treating malaria, or for raising antibodies for passive
 PT immunization.
 XX
 XX Claim 5; Fig 7; 73pp; English.
 XX
 XX The invention relates to new synthetic nucleotide sequences (I) encoding
 CC erythrocyte-binding protein (EBP) of a malaria pathogen, where codon
 CC usage of the synthetic nucleotide sequence is altered compared to a
 CC naturally occurring sequence of the EBP to approximate codon usage of a
 CC host of the malaria pathogen. (I) and the purified erythrocyte binding
 CC protein are useful for the production of a medicament for induction of an
 CC anti-malarial immune response in a human or animal. The methods are
 CC useful for inducing an anti-malarial immune response in a host of the
 CC malaria pathogen, optimizing expression in P. pastoris of a recombinant
 CC Plasmodium falciparum region II EBA-75 erythrocyte-binding protein, and
 CC obtaining the pure protein, and obtaining a DNA vaccine for inducing an
 CC anti-malarial immune response. (I) is useful in the production of anti-
 CC malaria DNA vaccines with improved immunogenicity, in expression vectors
 CC for the production of malaria EBPs, or as an immunization agent for
 CC inducing anti-malarial response and/or treatment of malaria. The EBPs are
 CC useful for raising antibodies or serum for passive immunization and
 CC treatment of malarial disease. The present sequence represents a malarial
 CC EBA-175 RII synthetic protein
 XX
 XX Sequence 616 AA;
 SQ
 Query Match 18.5%; Score 1199.5; DB 6; Length 616;
 Best Local Similarity 37.4%; Pred. No. 1.7e-66;
 Matches 252; Conservative 113; Mismatches 239; Indels 69; Gaps 15;
 121 NTNSNFPANTSIGKDNQYTFIQKTHLFCAGIKRKSITKICRENSKIT-VCPVDR 179
 3 NTSSN-----EV-----LSNCRKRGKMKWCKKNDKNSNTVCIPDR 40
 180 KIQLCVANFLNSRLTMEKFEIILISVNTBAKLLYNKNEGKPSIFCNELNSFSDFRS 239
 41 RIQLCIVNLSITKYTKETMDKHFIAESKSSQLLLKNDKNYSKFCNDLNSFLDVG 100
 240 SPIGDDMDPGNTRVKGYYINTKSDYYKKNVEKLNNIKKEWKEKKNANLWNIHVK 299
 101 LAMGNDMDFGYGYSTKAENKIQEVFGAHEISEHKIKNFRKKWNEFREKLWEAMLSHK 160
 300 GNISECAIIPABEPQINLWIKENENELMEKRLFLNIDKVCENKKEACFGCRLPC 359
 161 NNIN-NCKNIPQEBELQITQWIKWHGFEFLERDNRSLPKSKCNNTLYEACEKCIDPC 219
 360 SSYSFMMKSKTQMEVLNLYKKKNSGVDKNNFLDLFK-KNNKND--LDDFFKN-EKEY 415
 220 MKYRDWIIRSKFEWHTLSKEVETQK--VPKENAENYLIKSENKNDKVSLLNNCDAEY 277
 416 DDLCDCRYTATIKSFLNGPAKNDVDIASQINVDLRFQGNYS--NNEKSNWCTGTFT 473
 278 SKYCDCKRHTTLMKSVLNGNDNTIKEKREHIDLDDFSKFGCDKNSVDNTNTKVCCKPYK 337

Qy 474 NKFPGTCEPPRRQTLCLGRYTLHHRGHEEDYKSHLGLASIVEAQLLYKYKEDENALCS 533
 Db 338 LSTKDVCPVPRRQELCLGNIDRIYKLLMKEHILAIAYESRLKRYKNDKEVCV 397
 Qy 534 IIONSADLADIIGSDIIDYKYGKMBENLNKVKNDKKNNEESLKIIFREKWDENKENV 593
 Db 398 IINKTFADIRDIIGGTDYNDLSNRKLVGKINTNSYVHRNKQNDKLFDEWVKVKKOV 457
 Qy 594 WKMSAVLNKKECTCKDYDFKQIPQFLRWFKWGGDDFCCKRKEKIYSFESPKVECKKDC 653
 Db 458 WNVISWFKDKTVCKE-DDIENIPQFFRWFSEWGGDYCDQTKMI---ETLKVECKEPC 513
 Qy 654 DENTCKNKCSEYKKWIDLKXSEYKQV--DKYTKDNKKNMYNDIVKNEANVYLKKB 710
 Db 514 EDDNCKRKNCKSYKEWISKKKEEYNKQAKQYQYKGNMY---SEFKSIKPEVYLKCY 570
 Qy 711 SKECKDVNFDKIPNESPNEYEDMCKCDIKYLNEIKYPTKDDIYDIDTFSTFGDGT 770
 Db 571 SEKSNLNFDEBPKBELHSDYKKNCTWCPEVK-----DV 604
 Qy 771 PISINANINEQOS 783
 Db 605 PISIERN-NEQTS 616

RESULT 11
 AAM50533
 ID AAM50533 standard; protein; 616 AA.
 XX
 AC AAM50533;
 XX
 DT 12-MAR-2002 (first entry)
 XX
 DE Unidentified amino acid sequence.
 XX
 KW Merozoite surface antigen 1; MSP-1; p115MSP-1; antigen; immunogen;
 KW malaria; vaccine; Venezuela equine encephalitis virus; DVEE; vector;
 KW immunisation; Plasmodium falciparum.
 XX
 OS Unidentified.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 396 /note= "encoded by TCA"
 FT Misc-difference 512 /note= "encoded by CCT"
 FT
 XX WO200185927-A1.
 XX PN 15-NOV-2001.
 XX PD 08-MAY-2001; 2001WO-US014716.
 XX PF
 XX PR 08-MAY-2000; 2000US-0202430P.
 XX PA (GEOU) UNIV GEORGETOWN.
 XX
 FI Davidson E, Nikodem D;
 XX
 DR WPI; 2002-049444/06.
 DR N-PSDB; AAI70933.
 XX
 XX Novel vaccine for immunizing mammals against Plasmodium falciparum
 PT infection, comprises a viral vector system expressing protein
 PT corresponding to specific domain of major merozoite surface protein 1 of
 PT Plasmodium falciparum.
 XX
 PS Disclosure; Fig 6; 60pp; English.
 XX
 CC The present sequence is that of an amino acid DNA sequence given in
 CC Figure 6 of the invention. The identity of the sequence is not given. The
 CC invention relates to a malaria vaccine comprising an expression vector,
 CC preferably a defective Venezuelan equine encephalitis viral vector system,

Db 643 NNGSLQDQKTCQNVCTNMVYVTRKLAIEIQSVKY--DKRKLFP---SLAKDKNVTTF 697
Qy 707 LKEKSECKDVNPDDKIFNBNPEYEDMKKKD---EIKYLNIEKYPKTKHDIYDI--- 759
Db 698 LKENAKNCNIDF-TKIFDQDLKFKERCSCMDQVLEVKNEKMLSDSNSDSDATDISEK 756
Qy 760 -----DTPS 763
Db 757 NGBELVYVNHNSVSVASGNKEIEKSKDEKQPEKAKOTNGTLTVRTDKSDRNKKGDTAT 816
Qy 764 DT----- 765
Db 817 DTQNSPENLVQVHGHTNGETIKBPPKLPESSETLQSQEQLAEAAQKQKEEPPKKQEE 876
Qy 766 -----FGDGTPIISINANINEQOQSGK 785
Db 877 EPKKQBEQKQEQKQEQEERQKQEQEQIQQSQSGLDQSSKGVVASEQNEISSGQ 936
Qy 786 D-----TSNTGNETSDSPVSH-EPESDAAIN--VEKLSGDESSSETRGI 827
Db 937 EQNVKSSPEVVPQETTSNGSSQ--DTKISSTEPNENSVVDRATDSMNLDPKVNHNEM 994
Qy 828 LDIN-----DPSVTNNVNEVHDA-----SNTQ-----GSVSNTS 856
Db 995 SDPNTNTEPDASLKDCKEYDDAKKELQSTVSRIESNEQDVQSTPPEDTPTVEGKVDKA 1054
Qy 857 DI-----TNGHSESLN-----RTNN----- 872
Db 1055 EMLTSPHATDNSESESLNPTDDIKTTDGVVQEQEILGGESATETSKNLEKPKDVEPS 1114
Qy 873 ---AQDIKIGRSGNEQSD-----NOENSS-----HSDNSGSLTIGQ 906
Db 1115 HEISEPGLSVGTGKESELLKSKSIETKGTDPKSNQDEQDQDDVNSRDDNNLS--- 1171
Qy 907 VPSEDNTQNTYDSNP-----HRDTPNALASLPDDDKINIEGPDSS 948
Db 1172 -NSVDNQSNVLNREDPIASETEVVSPEPDSRSRIMTTEVPSTTVK-PDEKRSVEVGEKA 1229
Qy 949 RD-----SENGRG-DTTSNTHDVRRTNIVSER-----RVNSHDFIR 983
Db 1230 KEIKVEPVVPRAIQEPMEVSVQSPNVEDVEKETLISENGLHNDTHRGNISXKDLID 1289
Qy 984 NGMANNAHQYITQIENNG-IIRQESAGN-----SVNYKDNPKRSFSE----- 1030
Db 1290 IHLRNEAGSTILDSRRNGEMTEGSESDVGELQEHNFSTQQDKEDQDQIASDREKEEI 1349
Qy 1031 -----NDH-----KKNIQEYNSRDTKRVREIHK 1054
Db 1350 QKLNIGHDEDEDVLKMDRTEDSDGVSNSHLYNNLSSEKMEQYNNRDASKDREILN 1409
Qy 1055 LSKQKCNNEYSMEYCTYSDBERNSSPGCSREBERKKLCCQISDYCLKYFNFYIEYNCI 1114
Db 1410 RSNNTNCSNEHSLKYQYMERNKDILLETCSEDKRLHLCCEISDYCLKFFPKSIEYPDCT 1469
Qy 1115 KSEIKSPYKCFKSEGOSSPYFAAGGLVAVIVILLSSASRMGKSNEEYDIGESNI-EAT 1173
Db 1470 QKEPDDPTNCFKQKRTSMHYIAGGGIALLUFLILGSAVRKNLNDKDEKGFYDSNLDSA 1529
Qy 1174 FE-ENNYLNKLS-RIFNQEQVETNISDYSEYVNEKNMY 1210
Db 1530 FEYNNKYNKLPYMVVDQINNVNSDLYSEGIYDDTTTF 1568

RESULT 14

AAU76762
ID AAU76762 standard; protein; 1501 AA.
XX AAU76762;
AC AAU76762;
XX
XX
DT 21-MAY-2002 (first entry)
XX Plasmodium falciparum erythrocyte binding protein EBP5 protein sequence.
XX

KW Erythrocyte binding protein; EBP; malaria parasite; protozoacide;
KW vaccine; immune response inducer; Plasmodium falciparum merozoite; SABB;
XX salicylic acid binding protein; EBA-175.
OS Plasmodium falciparum.
XX
PN WO200211756-A2.
XX
XX 14-FEB-2002.
XX
XX 07-AUG-2001; 2001WO-US024725.
XX
XX 07-AUG-2000; 2000US-0223525P.
XX
XX (ENTR-) ENTREMED INC.
XX
XX Narum DL, Sim KL;
XX
XX WPI; 2002-227117/28.

Paralog of erythrocyte binding protein-175 polypeptide sequence useful
for inducing immune response to Plasmodium falciparum merozoites, that
causes malaria, in a patient, and for diagnostic and prognostic purposes.
Example 9; Fig 2; 52pp; English.

The present invention relates to a new polypeptide comprising a paralog
of erythrocyte binding protein-175 (EBA-175) polypeptide sequence. The
invention is useful for inducing an immune response to Plasmodium
falciparum merozoites in a patient. The method of the invention comprises
administration of isolated SABB (salicylic acid binding protein) binding
domain polypeptide. The invention is further useful for research
applications, vaccine candidates, blocking peptides, diagnostics and
prognostics, where diseases of Plasmodium falciparum include malaria. The
polypeptide of the invention and antibodies are useful for diagnosis of
malaria, and for detection of P. falciparum in culture media and in
biological samples such as biological tissues and fluids. The present
amino acid sequence represents the Plasmodium falciparum EBP5 protein
which is one of several (AAU76759-AAU75762) EBA-175 paralogues
Sequence 1501 AA;

Query Match 15.3%; Score 993.5; DB 5; Length 1501;
Best Local Similarity 21.7%; Pred. No. 5.1e-53;
Matches 343; Conservative 228; Mismatches 472; Indels 537; Gaps 63;

Qy 1 MKGVFNI-----YFLIPLIFLYNVIRINESIIGRTLYNRQDESSDISRVNSPELNNHKT 55
Db 3 MKGKMMKCLFFFYIILYVVLCTYVLGISEYL-----KERPQGLNVETNNNNNNNNNS 57
Qy 56 NIYDSYEDVNNKLINSFVEN-KSVKKKRSLSFINNKTYSYDIIPPSYSYRNDKFNLSLE 114
Db 58 N--SNDAMSFVNEVIR-FIENEKDDKEDKKVKIISRPVEN-----TLHRYVPVSSFLNIKK 109
Qy 115 NEDNSGNTSNPNFANTSEISIGKDNKQYTFIQKTHLFCAGIKRKRKIKWICRENSKITV 174
Db 110 YGRKGTYLARNSPV-----QSRVIRGCKGRKSTHTWIC-ENKGNNI 150
Qy 175 CVPDRKIQLCVA--NFLNSRLTMEK--FKEIFLSVNTAEKLLYNKNGKDPSPICNE 229
Db 151 CIPDRVQLCITALQDLKNSGSETTDKRLLRKVFDSAMYETOLLWNKYGFRDFDPCDD 210
Qy 230 LRNSFDSFRSSFIGDDMDFGNTDRVKGYINTKFSYDKYKKNVEKNLNNIKKEWKNKAN 289
Db 211 VKNSYLDYKDVIFGTOLD---KNNISKLVEESLKRFFK-KDSSVLN--PTAWRRYGR 263
Qy 290 LWNHMI--VNHKGNISKECAIIPAEPPQINLWKEH---NENFLMEKKRLFLNIKKQCV 344
Db 264 LMKWTMQPYAHLG-----CRKPDENEPPNRIWLEWGYKNCRLMKEKEKL---LTGEC 315
Qy 345 NKKYEACFGGCRLLPCSSYTSFMKSKTOMEVLTNLVKK-----KNSGVYDKNNFLNDF 397
Db 316 NRKSDCSGTCNNECTYRSNLNRQYEVSIILGKKIKVRYTIFRKRIVQPNAL-DFL 374

QY 398 KNNKNN-----DLDDFFKNE-KEYDDLCRCRYTATIIKSLFNGPAKNDVDIASQINVDLR 452
Db 375 KLNCECKDIDFKFFBFEYKYEKMC-----QSYIDLKIQKNDI- 418
QY 453 GFGCNYSKNE-----KSWNCTGFTTNKF-----PGTCEPPRRQTLCLGR 492
Db 419 ---CSFNAQTIVSSDKRFLCKEKEFKWKCD---KNSFETVHHKGVCSPPRQGFCLGN 472
QY 493 -TYLLHRRHEDEYKEH-----LLGASIVBAQLLKYYKE-KDENALCSIIQNSYADLA 543
Db 473 LNYLL---NDIYNVHNSQLIBIIMASKQKGLKWKKGHTILDNQACKYINDSYDYK 529
QY 544 DIIRGSDIIKDYKGMKEENL-----KVNKKRNEESLKIIFREKWDENKENV 594
Db 530 DIVIGNDLNDNNSIKVQNNLNLIIFERNFYKVGNKL--FKTIKELKNVWILNRNVW 587
QY 595 KVM-----SAVLNKETCKDYKFKIPOFLRWFKEWDDPCEKKEKIKYISPEKVECK- 649
Db 598 ESMRCGIDEVDQRKTCERIDELENMPOFFRWFQWAFHFC---KEKEYWELKLNCKCTG 644
QY 650 ---KDCDENTCKNCSYKWKIDILKSEYKQVDKYTKDKNKKMYDNIDEVKNKEANVY 706
Db 645 NNGSLQODTKQNVCTNMNTWTYTKLAYBIQVKY--DKDKLF---SLAKDNVTFP 699
QY 707 LKESKECKOVNDDKIFNESPNBYEDMCKCD---EIKYLNEIKYKPKTKHDIYDI--- 759
Db 700 LKENAKNSNIDF-TKIPDQLDKLPERCSMDQVLEKVKEMLSIDNSSEDATDISEK 758
QY 760 -----DTFS 763
Db 759 NGBEELYNNHNSVSVASGNKEIEKSDEKQPEKAQTNGTLTVRTDKSDRNKGDTAT 818
QY 764 DT----- 765
Db 819 DTKNSPENLKVQEHGTNGETIKERBPPLPESSETLQSOQLAEBAQKQKQEBEPKQKE 878
QY 766 -----FGDGTPIISANININEQOSGK 785
Db 879 EPKKQEBEQKQEQKQEBEQKQEQEQIQOSQSGLDQSKVGVASEQNEISGQ 938
QY 786 D-----TSNTGNSSTSPVSH-BPESDAAN--VEKLSGDESSETRGI 827
Db 939 EQNVKSSSPVPOETTSENGSSQ--DTKISSTEPNNSVVDRTDSMNLDPKPVHNM 996
QY 828 LDIN-----DPSVTNNVNEVDA-----SNTQ-----GSVGNYS 856
Db 997 SDPNTNTEPDASLKDDKKEVDDAKKELQSTVSRIESNEQDVQSTPPEDTPTVEKVGDKA 1056
QY 857 DI-----TNGHSESSLN-----RTTN----- 872
Db 1057 EMLTSPHATDNSSESGLNPTDDIKTDGVVKEQEIILGGESATETSKSNLEKPKDVEPS 1116
QY 873 ---AQDKIGRSGNEQSD-----NQENSS-----HSDNSGSLTIGQ 906
Db 1117 HEISPEVLGTTGKESELLKSKIETKGETDPRNSDQEDATDDVENSRRDNNLS--- 1173
QY 907 VPSEDNTQNTYDSQP-----HRTPNALASLPDDKINEIEGFDSS 948
Db 1174 -NSVDNQSNNLNRDPDIASETEVSVSEPDSSRIITTEVPSTTVK-PPDEKRSSEVGEKA 1231
QY 949 RD-----SENGRG-DTTSNTHDVRNTIVSER-----RVNSHPFIR 983
Db 1232 KEIKVEPVVPAIGEPNNSVSVGPPNVEDVEKETLISENGLNTHRGNISSEKULID 1291
QY 984 NGMANNNAHQYITQIENNG-IIRQEBESAGN-----SVNYKDNKPRNFSSE----- 1030
Db 1292 IHLRNEAGSTILDDSRNGEMTGESDVGELQEHNFSTQKQKEDQDQIASDREKEE 1351
QY 1031 -----NDH-----KKNIQEVNSRDTKRVREHIIK 1054
Db 1352 QKLLNIGHEDEVDLKMDRTESDMSDGVNSHLYNNLSSEKMEQYNNRASKDREELN 1411

QY 1055 LSKONKCNNEYSMEYCTVSDERNSSPGCSREERKCLCCQISDYCLKYFNFSYIEYVNCI 1114
Db 1412 RSNVTNTCSNEHSLXKCYQYMERKNDLLTSCSDKXRLHLCCEISDYCLKFPKPSIEYFDCT 1471
QY 1115 KSEIKSPYCKFKSEGGQSSI 1134
Db 1472 QKEFDDPTNGFRKQRTSM 1491
RESULT 15
AARI3457
ID AARI3457 standard; protein; 1115 AA.
XX
AC AARI3457;
DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 17-DEC-2001 (revised)
DT 07-NOV-1991 (first entry)
XX
DE Duffy receptor.
XX
KW Malaria; parasite; vaccine.
XX
OS Plasmodium vivax; Salvador I strain.
XX
FH Key Location/Qualifiers
FT Region 1..1050
FT Peptide 1..22
FT /label= exon 1
FT /label= signal sequence
FT Region 1051..1077
FT /label= exon 2
FT Region 1078..1101
FT /label= exon 3
FT Region 1102..1115
FT /label= exon 4
XX
USN7554837-N.
XX
PD 23-JUL-1991.
XX
PF 20-JUL-1990; 90US-00554567.
XX
PR 20-JUL-1990; 90US-00554837.
XX
PA (USSH) NAT INST OF HEALTH.
PA (USDC) US SEC OF COMMERCE.
XX
PI Miller L, Adams J, Kaslow D, Xiangdong P;
XX
XX
XX WPI; 1991-260184/35.
XX N-PSDB; AAQ13317.
XX
XX Plasmodium Duffy receptor proteins - used in vaccines against malaria and
XX for producing antibodies which prevent malaria infection.
XX
XX Disclosure; Fig 12; 63pp; English.
XX
XX The sequence was deduced from clones isolated from a P. vivax genomic
XX library. The sequence covers 4 exons and has a 22 amino acid signal
XX sequence, an 18 AA transmembrane segment (exact posn. not given) followed
XX by 45 AAs at the C-terminus. The protein lacks the repeat regions of the
XX corresponding protein from P. knowlesi but both have two regions of high
XX Cys content are separated by a pro- rich region. The gene can be used to
XX express recombinant Duffy receptor proteins for use in vaccines for
XX malaria. See also AARI3456. (Note: Revised entry submitted to correct the
XX patent number format of US Government-owned NTIS applications to prevent
XX clashes with ongoing US granted patent numbers. For further information
XX please visit the Derwent web site at
XX www.derwent.com/dwpi/updates/ntis us.html.) (Updated on 25-MAR-2003 to
XX correct PR field.) (Updated on 25-MAR-2003 to correct PA field.) (Updated
XX on 24-OCT-2003 to standardise OS field)

Db 30 CRSQPEFNK-----IPERNVQI-----HISNIFKEYKENNVDIIFGTLNLYEYNN 74
 Qy 418 LC-DCRYTATTIIFSLNGPAKNDVDIASQINVDNLGRFGCNYK--SNNEKSNWC--TGFTFN 474
 Db 75 FCKEPELVAANKYLNKAPNAKSPIRYKSEHESSVFGCKTKISKYKKNWCYNNKVT 134
 Qy 475 KPGTCPPRRQTLCLGRTYLLHRGHEEDYKHLGLGASIIYEAQLLKYYKEKDNALCSI 534
 Db 135 KPEGVCGPPRRQQLCLGYIFLIRDGNEBGLKDHINKAANYEAMHLKEKYENAGDKICNA 194
 Qy 535 IONSADLADIIGSDIIKSDYKGMENLNK-----NKDKRNEESLKIIFREKWDEN 589
 Db 195 ILGSYADIGDIVRGDLVWRDINTNKLSEKFKIFMGGSNGSKKQNDNN---ERNKWNKQ 251
 Qy 590 KENVKWSAVLNKKECKDYDFKQIPQFLRWFKEWGGDFCFKREKIKIYSEFVKVECK 649
 Db 252 RNLIWSMVKHIPKGTCKRHNNFEKIPQFLRWLKEWGGDFCEEMGTVEVKOLEKI---CE 308
 Qy 590 KENVKWSAVLNKKECKDYDFKQIPQFLRWFKEWGGDFCFKREKIKIYSEFVKVECK 649
 Db 252 RNLIWSMVKHIPKGTCKRHNNFEKIPQFLRWLKEWGGDFCEEMGTVEVKOLEKI---CE 308
 Qy 650 KDCDENTCKNCKSEYKQKWDILKSEYKQVDTYKOK--NKK--MYDNDIDEVNKEANV 705
 Db 309 NKNCSEKCKNACSSVEKWKERKNEYNLQSKFDSKLNKNNLYNKPEDSK-----A 363
 Qy 706 YLKEKSECKDVNPDKKIFNESPNEYEDMCKCD---EIKYLNKIK---YPKTKHDIYDI 759
 Db 364 YLRSKQCSNIEFNDETFT--FPNKYKEACMVCENPSSSKALKPKIKTNVFPBESKSEL 422
 Qy 760 DTFSDTFGDTPIISINANEQSGK--DTSNTGNSETSDSPVSHPESDAAINV-----812
 Db 423 SSLTDSKNTNSGGNGYGRQISKRDDVHHDGPKVK--SGEKEVPKIDAAVKTENEFT 481
 Qy 813 -----EKLSDGESSETRGILDINDPS--VTNNVNEVHDASNTQGSVNSTDITNG 861
 Db 482 SNRNDIEGKSKGDSHPVSHSKDKNEEPORVVSLENLPKIEKQWESSDSIPIT-----535
 Qy 862 HSSESLARTNAQD-----IKIGSGNEQS-----DNQENSHSDNSGLTIGQVPS 909
 Db 536 HIEAKQGSNSNDPAVSGRESKQVNLHTSRIKENEGVITKDDSSKSIKIPS 595
 Qy 910 EDN-----TQNTYDSQNHRTDTPNALASLPDDK--INEIEGFDSSRDSENGRDTTSNT 962
 Db 596 DQNNHSDLSQANEDSQNKETINP-----PSTEKNLKEIHYKTSDDSDHSGSKISIEP 651
 Qy 963 HDVARTNIVSERR-----VNSHDFIRNGMANNNAHHQYITQIENNGIIRGQE 1009
 Db 652 KELTEBSPLTDKTESAAIGDKNHSVKSADIFQSEIHNSDNRDRIVSE-----SVW---Q 704
 Qy 1010 ESAGNSVNYK-----DNPKRNSFSSNDHKKNIQBYNSR 1043
 Db 705 DSSGSSMTSEIRTDN---KDFKTSEDIAPSINGRNSR 739

RESULT 17
 AAW22479
 ID AAW22479 standard; protein; 749 AA.
 AC AAW22479;
 XX AAW22479;
 DT 07-OCT-1997 (first entry)
 XX Plasmodium ebl-1.
 DE DBL gene family; SAPP; sialic acid binding protein; vaccine; therapy;
 KW Duffy binding like gene; Duffy antigen binding protein; erythrocyte;
 KW DAPP; merozoite; malaria; var-1; var-2; var-3; var-7; immune response;
 KW Plasmodium.
 XX Plasmodium falciparum.
 OS WO9640766-A2.
 XX 19-DEC-1996.
 XX 07-JUN-1996; 96WO-US009508.

XX 07-JUN-1995; 95US-00487826.
 PR (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA Sim KL, Chitnis C, Miller LH, Peterson DS, Su X, Wellems TE;
 FI WPI; 1997-052231/05.
 XX N-PSDB; AAT72894.
 PT New malaria vaccines - contains cysteine-rich DBL family protein binding
 XX domains homologous domains of the Duffy and sialic acid binding proteins.
 PS Disclosure; Page 41-42; 96pp; English.
 CC This sequence represents ebl-1 of Plasmodium. Ebl-1 belongs to the Duffy
 CC binding like (DBL) family of genes which have homology to the Duffy
 CC antigen binding protein (DAPP) and sialic acid binding protein (SAPP)
 CC conserved regions (see AAT72889 and AAT72888 respectively). The var
 CC family of genes modulate cytoadherence and antigenic variation of
 CC Plasmodium infected erythrocytes. SAPP and the Duffy antigen binding
 CC protein (DAPP) are soluble proteins that appear in the culture
 CC supernatant after infected erythrocytes release merozoites. DAPP and SAPP
 CC mediate the binding of merozoites and schizonts to the erythrocyte
 CC surface. These proteins are necessary for erythrocyte invasion by the
 CC parasite. This sequence can be used in the compositions of the invention.
 CC The compositions are for the treatment and prevention of malaria, and
 CC comprise either a nucleotide sequence or encoded polypeptide of the var-
 CC 1, var-2, var-3 or var-7 genes of the DBL gene family, a family of genes
 CC having homology with conserved regions of DAPP and SAPP. The compositions
 CC are used for the treatment and prevention of malaria. They are also used
 CC in the preparation of vaccines for inducing a protective immune response
 CC in a mammal to Plasmodium merozoites (especially Plasmodium falciparum or
 CC Plasmodium vivax)
 XX Sequence 749 AA;

Query Match 11.7%; Score 761.5; DB 2; Length 749;
 Best Local Similarity 27.3%; Pred. No. 7.2e-39;
 Matches 207; Conservative 148; Mismatches 282; Indels 121; Gaps 31;
 Qy 359 CSSYTSFMKSKTQMEVLNLYKKNKSGVDKNNFLNDFKNNKNDLDDFPKN-EKEYDD 417
 Db 30 CRSQPEFNK-----IPERNVQI-----HISNIFKEYKENNVDIIFGTLNLYEYNN 74
 Qy 418 LC-DCRYTATTIIFSLNGPAKNDVDIASQINVDNLGRFGCNYK--SNNEKSNWC--TGFTFN 474
 Db 75 FCKEPELVAANKYLNKAPNAKSPIRYKSEHESSVFGCKTKISKYKKNWCYNNKVT 134
 Qy 475 KPGTCPPRRQTLCLGRTYLLHRGHEEDYKHLGLGASIIYEAQLLKYYKEKDNALCSI 534
 Db 135 KPEGVCGPPRRQQLCLGYIFLIRDGNEBGLKDHINKAANYEAMHLKEKYENAGDKICNA 194
 Qy 535 IONSADLADIIGSDIIKSDYKGMENLNK-----NKDKRNEESLKIIFREKWDEN 589
 Db 195 ILGSYADIGDIVRGDLVWRDINTNKLSEKFKIFMGGSNGSKKQNDNN---ERNKWNKQ 251
 Qy 590 KENVKWSAVLNKKECKDYDFKQIPQFLRWFKEWGGDFCFKREKIKIYSEFVKVECK 649
 Db 252 RNLIWSMVKHIPKGTCKRHNNFEKIPQFLRWLKEWGGDFCEEMGTVEVKOLEKI---CE 308
 Qy 650 KDCDENTCKNCKSEYKQKWDILKSEYKQVDTYKOK--NKK--MYDNDIDEVNKEANV 705
 Db 309 NKNCSEKCKNACSSVEKWKERKNEYNLQSKFDSKLNKNNLYNKPEDSK-----A 363
 Qy 706 YLKEKSECKDVNPDKKIFNESPNEYEDMCKCD---EIKYLNKIK---YPKTKHDIYDI 759
 Db 364 YLRSKQCSNIEFNDETFT--FPNKYKEACMVCENPSSSKALKPKIKTNVFPBESKSEL 422
 Qy 760 DTFSDTFGDTPIISINANEQSGK--DTSNTGNSETSDSPVSHPESDAAINV-----812
 Db 423 SSLTDSKNTNSGGNGYGRQISKRDDVHHDGPKVK--SGEKEVPKIDAAVKTENEFT 481

QY 813 -----EKLSDGESSSTRGILDINDPS--VTNNVNEVHDASNTQGSVNTSDITNG 861
 Db 482 SNRNDIEGKESKGDHSPVHSKDINKKEPQVVSENLPIKEKMSDSDIPIT----- 535
 QY 862 HSESSLNRTTNAQD-----IKIGSGNEQS-----DNOENSHSDSGSLTIGOVPS 909
 Db 536 HIEAEKGQSSNSDNDPAVVGRESKDVNLHTSERIKENEVGIKTDSSKSIEKIPS 595
 QY 910 EDN-----TON-TYDSQNPHRDTPNALASIPSDK-INEIEGFDSSRDSNGRDTTSNT 962
 Db 596 DQNNHSDLSQANEDSNQGNKETINP-----PSTEKMLKEIHYKTSDDSDHGSIKSIEP 651
 QY 963 HDVVRTNIVSERR-----VNSHDFIRNGMANNNAHQQYITQIENNGIRGOE 1009
 Db 652 KELTESPLTDKKTESAAIGDKNHESVKSADIFQSEIHNSDNRDRIVSE-----SVV---Q 704
 QY 1010 ESAGNSVNYK-----DNPGRSNFSSNDHKNIQYNSR 1043
 Db 705 DSSGSSMSTESIRTDN---KDFKTSEDIAPSINGRNSR 739
 RESULT 18
 AAY77901
 ID AAY77901 standard; protein; 749 AA.
 AC AAY77901;
 XX
 DT 13-JUN-2000 (first entry)
 XX
 DE P. falciparum ebl-1 polypeptide.
 XX
 KW DBL gene; Duffy-binding like gene; ebl-1; Duffy Antigen Binding Protein;
 KW DABP; Sialic Acid Binding Protein; SABP; malaria; vaccine; immunisation;
 KW protozoacide.
 XX
 OS Plasmodium falciparum.
 XX
 FN US5993827-A.
 XX
 PD 30-NOV-1999.
 XX
 PF 07-JUN-1995; 95US-00487826.
 XX
 PR 10-SEP-1993; 93US-00119677.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Sim KL, Chitnis C, Peterson DS, Su X, Wellem's TE, Miller LH;
 XX
 DR WPI; 2000-194198/17.
 DR N-PSDB; AAZ98283.
 XX
 PT Isolated protein binding domains from Plasmodium vivax and Plasmodium
 PT falciparum erythrocyte binding proteins useful for vaccinating against
 PT malaria.
 XX
 PS Claim 7; Col 53-58; 93pp; English.
 XX
 CC The invention relates to ebl-1 polypeptides that are encoded by the DBL
 CC (Duffy-binding like) gene family. The ebl-1 proteins are substantially
 CC identical to the Duffy Antigen Binding Protein (DABP) and sialic acid
 CC Binding Protein (SABP), which are soluble proteins that appear in the
 CC culture supernatant after erythrocytes infected with malaria release
 CC merozoites. Immunohistochemical studies indicate that DABP and SABP are the
 CC respective ligands for Plasmodium vivax and Plasmodium falciparum Duffy
 CC and sialic acid receptors on erythrocytes. The ebl-1 polypeptides may be
 CC used to vaccinate against malaria, especially caused by P. falciparum.
 CC Immunization with the polypeptide provides effective protection against
 CC malaria. The present sequence represents the ebl-1 polypeptide
 XX
 SQ Sequence 749 AA;

Query Match 11.7%; Score 761.5; DB 3; Length 749;

Best Local Similarity 27.3%; Pred. No. 7.2e-39;
 Matches 207; Conservative 148; Mismatches 282; Indels 121; Gaps 31;
 QY 359 CSSTYTFMKSKTKQMEVLNLYKKKSGVDKNNFLNDFKNNKNDLDDPFKN-EKEYDD 417
 Db 30 CRSGPFPK-----IFPERVQI-----HISNIFKEYKENVDIIFGLTNYNN 74
 QY 418 LC-DCRYTATIIKSFNGPAKNDVDIASQINVDLRFQGCNYK-SNNEKSWNC-TGTFTN 474
 Db 75 FCKEPELVSAAYNLKAPNAKSPRIYKSEHESSVFGCKTKISKVKKKWCYNNKVT 134
 QY 475 KPFGTCPEPRRQTLCLGRYVLLHRGHEEDYKEHLLGASTYEAQLLYKYKKEKENALCSI 534
 Db 135 KPEGVCGPFRROQLCLGYIFLRDGNBGLKDHINKAANYEAMHLKEKEYNAGDKTCNA 194
 QY 535 IQNSYADLADIIGSDIIOYGYKMEENLKV-----NKDKRNEESLKIPEKMWDEN 589
 Db 195 ILGSYADIGDIVRGLDVWRDINTNKLSEKFKIPMGGSRSRKKQNDNN---ERNKWEKQ 251
 QY 590 KENVKMSAVLKNKETCKDYDKFKIPQFLRWFKWGDGDFCKRKEKIYSPESFKVECK 649
 Db 252 RNLWSSMVKHIPKGTCKRHNFXEIPQFLRWLKEWGDGDFCEMGTEVKOLEKI---CE 308
 QY 650 KKDCCENTCKKCYKXKIDLKSEYKQVDKYTKK--NKK--MYDNIDEVKNKEANV 705
 Db 309 NKNCSKKCKKACSSYERIKERKNYNLQSKKFDSDKLNKNLYNKFEDSK-----A 363
 QY 706 YLKEKSKCKDVNFDKIFNESNEYEDMCKKCD---EIKYLNKIK---YPKTKHDIYDI 759
 Db 364 YLRESKQCSNIFNDETF-T-FNKYKEACWVCENPSSSKALPKITNVPPIESKSEL 422
 QY 760 DTFSDTFGDTPIISANINEQSGK--DTSNTGNSSETSDSPVSHBPESDAAINV----- 812
 Db 423 SSLTDKSKNTPNSSGGNGYDQRIQSKRDVHDGPKVK--SGEKEVPKIDAAVKTENEFT 481
 QY 813 -----EKLSDGESSSTRGILDINDPS--VTNNVNEVHDASNTQGSVNTSDITNG 861
 Db 482 SNRNDIEGKESKGDHSPVHSKDINKKEPQVVSENLPIKEKMSDSDIPIT----- 535
 QY 862 HSESSLNRTTNAQD-----IKIGSGNEQS-----DNOENSHSDSGSLTIGOVPS 909
 Db 536 HIEAEKGQSSNSDNDPAVVGRESKDVNLHTSERIKENEVGIKTDSSKSIEKIPS 595
 QY 910 EDN-----TON-TYDSQNPHRDTPNALASIPSDK-INEIEGFDSSRDSNGRDTTSNT 962
 Db 596 DQNNHSDLSQANEDSNQGNKETINP-----PSTEKMLKEIHYKTSDDSDHGSIKSIEP 651
 QY 963 HDVVRTNIVSERR-----VNSHDFIRNGMANNNAHQQYITQIENNGIRGOE 1009
 Db 652 KELTESPLTDKKTESAAIGDKNHESVKSADIFQSEIHNSDNRDRIVSE-----SVV---Q 704
 QY 1010 ESAGNSVNYK-----DNPGRSNFSSNDHKNIQYNSR 1043
 Db 705 DSSGSSMSTESIRTDN---KDFKTSEDIAPSINGRNSR 739
 RESULT 19
 AAW22478
 ID AAW22478 standard; protein; 1115 AA.
 XX
 AC AAW22478;
 XX
 DT 12-SEP-1997 (first entry)
 XX
 DE Duffy antigen binding protein.
 XX
 KW DBL gene family; SABP; sialic acid binding protein; merozoite; malaria;
 KW Duffy antigen binding protein; DABP; erythrocyte; var-1; var-2; var-3;
 KW var-7; vaccine; therapy; immune response; Plasmodium.
 XX
 OS Plasmodium vivax.
 XX
 PH Key Location/Qualifiers

FT Misc-difference 1065. .1067
 TT /note= "encoded by CTGCTGTGTTA"

PN W09640766-A2.

PD 19-DEC-1996.

XX 07-JUN-1996; 96WO-U009508.

XX 07-JUN-1995; 95US-00487826.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Sim KL, Chitnis C, Miller LH, Peterson DS, Su X, Wellens TE;

XX WPI; 1997-052231/05.

XX N-PSDB; AAT72889.

XX New malaria vaccines - contains cysteine-rich DBL family protein binding domains homologous domains of the Duffy and sialic acid binding proteins.

XX Example 1; Page 33-35; 96pp; English.

XX This sequence represents the full length Duffy antigen binding protein (DABP). DABP and the sialic acid binding protein (SABP) are soluble proteins that appear in the culture supernatant after infected erythrocytes release merozoites. DABP and SABP mediate the binding of merozoites and schizonts to the erythrocyte surface. These proteins are necessary for erythrocyte invasion by the parasite. This sequence can be used in the compositions of the invention. The compositions are for the treatment and prevention of malaria, and comprise either a nucleotide sequence or encoded polypeptide of the var-1, var-2, var-3 or var-7 genes of the DBL gene family, a family of genes having homology with conserved regions of DABP and SABP. The compositions are used for the treatment and prevention of malaria. They are also used in the preparation of vaccines for inducing a protective immune response in a mammal to Plasmodium merozoites (especially Plasmodium falciparum or Plasmodium vivax)

XX Sequence 1115 AA;

Query Match 11.6%; Score 754; DB 2; Length 1115;

Best Local Similarity 24.5%; Pred. No. 3.7e-38;

Matches 296; Conservative 177; Mismatches 416; Indels 318; Gaps 57;

QY 199 FKEIFLSVNTAEKLLYNKNEGDPSTFCNELRNSFSDFRSSFFGDDWDFGNTDVRKYG 258

DB 22 YKDDFSIT-----LNYHEGKKYLILKR-----KLEKANNRDVNCFF 59

QY 259 IN-TKFSDYKKNVEKLNINIKKEWBE-KKANLWNNHIVNHKGNISKECAIIPAEFPQI 316

DB 60 LHPFQVNNVLLERTIETLLCEKNEYVKGNGYKL-----AKGH---HC-----VEEDNL 105

QY 317 NLWIKENWNPFLMKRLFLNIKDKVENKYEACFGGRLPCS-SYTSFPMKSKTOMEV 375

DB 106 ERWLOQTNER-----RSEENIKYK--YGVTELKIKYQAQNGKRSSRIILKES 149

QY 376 L-----TNLYKKKSGVDKNNFLNDLPKNNKNDLDDFPKNEKYEDDLCDCHY----- 423

DB 150 IYGANFGNSYMEKGDDGDKTG-----EEKDGEHKTDSKTDNGKANNVLMLDYETSSN 204

QY 424 -----TATIIFKSLFNGPAKN-----DVD-----IASQINVNDLGRF-----G 455

DB 205 QOPAGTLNDLVLEFVTGHEGNSRKSSNGNPYDIDHKTTISSAI-IN-HAFLQNTWKN 261

QY 456 CNYK-SNNKSWNCTGFTTNKPGTCPPRQTCLGRTVLLHRGHEEDY-----KEH 507

DB 262 CNYKRRERDWC-----NTKQDVCIPDRRYQLCMKELTNVNTDTHFRDITPRKLY 316

QY 508 LLGASIIYA-----QLLYKYKEKDEKALCSIIQNSVADLADIKGSIDIKYGKK 559

DB 317 LKRRLIYDAAVEGDLKLKANNRY-NKD---FCKDIRWSLGDGFDIIMGTMEGIGYSKV 372

QY 560 MEENLNKV-NKDKKRNEESLKI PREKWNDEKENKENVKMSAVLNK-----ETCKDYDK 612

DB 373 VENNLSIFGTDEKAQR-----RKQWNSSQAQIWTAMVSVKRLKGNFIWICKLVA 427

QY 613 FQKIPQELRWFKEWDDFCRKEKEKYSFSPKVC-----KKQDCDENTCKNKSEY 665

DB 428 VNTPEQIYRIWREGRDYSELPTFV---OKLKEKCDGKINYTDKVKYKVPCCNACKSY 484

QY 666 KKWIDLKSEYEKQVDKYTKDKNKNMYDNIDEVKNKENV-----YLKESKCKECKVNF 720

DB 485 DQWITRKNQWDLVSNKFISVKNAB-----KVQTAGIVTPVDYILKQBLDEFENVAPE 536

QY 721 DKTFNSPNYEDMCKKCDKBEIKYLNEIKYKPKTHDIYDIDTFSDTFG-DGTPI----- 772

DB 537 NEI-NKRDGAYIELC-----VCSVEEAK-KNTQEVVTVDNAAKSOATNSNPISQPVDS 589

QY 773 -----SINANINEQSGKDTNTCNSETSP-----SPVSH-----PESDAAINV 812

DB 590 KAEKVPGDSTHGNNV---SQDSSITTKGAVTGGQNGNTPAESDVQORSIDIAESVSANV 646

QY 813 E-----KLSGD-----ESSSETRGILDINDPSPVTNNVNEVHDASNTQGSV----- 852

DB 647 DPQKSVSKRSDDTASVTGIAEACKENLGASNSRPSSESTVEANSFGDDTVNSASIPVVS 706

QY 853 -----SNTSDITNGHSESSLNRTTNAQDIKIRSG-----NEQSDNQNSHSS 896

DB 707 NPLVTPYNGLRHKSNDSDGPAESMANPDSNGK---GETGKQDNDMAKATKDSNS 762

QY 897 DNGSLT-----IGQVSEDTQNTQYDSQNPHRDTPNALASLPSDDKINEIEG 944

DB 763 DGTSSATGDTTDAVDREINKGVPEDRDKTVGSKDGGEDNSANKDAATVVGEDRRENSA 822

QY 945 FDSRD-----SENG-----RGDTTSN---THDVRRTNIV 971

DB 823 GGSNTDRSKNDTEKNGASTPDSKQSEDATAKLSKTESLESTESGDRITNDTNSLEKNG 882

QY 972 SERRVNSHDFIRGMANNNAHQYIT-----QIENNGIIRQSEESAGNSVNYKDNPK 1023

DB 883 KERDLQKHDFKSNPTNEBPSDQTTDAEGHDRDSIKND---KAERKHKMKDFTTKNTN 939

QY 1024 RSNFSEN---DHKKNIQEYNSRDTKRVREIILKSKQKCNKNEYSNEYCTYSDERNSSP 1080

DB 940 SHLNSNNLNSGKLDIKEYKYRDVATREDIILMSSVRKCNNNISLEYCNVEDKISS- 998

QY 1081 GPCSRERKKLCCOISDYCLKYFNFYSLIYVNCIKSEIKSPEYKCPKSEG-QSSIPYFAA 1139

DB 999 NTSRKSRLKCCSISDFCLNYFDVYSYELVSKKEFEPSYKCTKGKFKIDKTYFAA 1058

QY 1140 GGLVWIVILLSSASRMKSNBEYDIGESNIEATPEE-NNYLNKLSRI-----FNQEVQE 1193

DB 1059 AGAL--LILLILLASRMKKNDS-----EATFNEFEYCDNIHRIPLMPNNIEHQ 1108

QY 1194 TNISDYS 1200

DB 1109 STPLDYS 1115

RESULT 20

AAY77899

ID AAY77899 standard; protein; 1115 AA.

XX AC AAY77899;

XX AC AC

DT 13-JUN-2000 (first entry)

XX DE DBL gene; Duffy-binding domain polypeptide.

XX DE DBL gene; Duffy-binding like gene; ebl-1; Duffy Antigen Binding Protein;

KW DABP; Sialic Acid Binding Protein; SABP; malaria; vaccine; immunisation;

KW protozoa.

XX OS Plasmodium vivax.

XX PN US5993827-A.

30-NOV-1999.
07-JUN-1995; 95US-00487826.
10-SEP-1993; 93US-00119677.
(USSH) US DEPT HEALTH & HUMAN SERVICES.
Sim KL, Chitnis C, Peterson DS, Su X, Wellem's TE, Miller LH;
WPI; 2000-194198/17.
N-PSDB; AAZ98281.
Isolated protein binding domains from Plasmodium vivax and Plasmodium falciparum erythrocyte binding proteins useful for vaccinating against malaria.
Example 1; Col 35-40; 93pp; English.
The invention relates to ebl-1 polypeptides that are encoded by the DBL (Duffy-binding like) gene family. The ebl-1 proteins are substantially identical to the Duffy Antigen Binding Protein (DABP) and Sialic Acid Binding Protein (SABP), which are soluble proteins that appear in the culture supernatant after erythrocytes infected with malaria release merozoites. Immunochemical studies indicate that DABP and SABP are the respective ligands for Plasmodium vivax and Plasmodium falciparum Duffy and sialic acid receptors on erythrocytes. The ebl-1 polypeptides may be used to vaccinate against malaria, especially caused by P. falciparum. Immunization with the polypeptide provides effective protection against malaria. The present sequence represents the DABP binding domain polypeptide
Query Match 11.6%; Score 754; DB 3; Length 1115;
Best Local Similarity 24.5%; Pred. No. 3.7e-38;
Matches 296; Conservative 177; Mismatches 416; Indels 318; Gaps 57;
QY 199 FKEIFLSVNTAEKLYNKGKDPISFCNELNSFSDPRSSFTGDDMDFGGNTDRVKY 258
DB 22 YKDFISIT-----LYNHEGKKYLILKR-----KLEKANNRDVCNFF 59
QY 259 IN-TKFSDDYKKNVKNLNIKKWBE-KKANLWNHMI VNHKNISKECAIIPAEBPQI 316
DB 60 LHFSQVNNVLLERTIETLLECKNEYVKGNGYKL-----AKGH-----HC-----VEEDNL 105
QY 317 NLWIKENWENFLMEKKRLFLNIKDCKVENKYEACFGGCRLPSCS-SYTSFMKSKTQMEV 375
DB 106 ERWLQGTNER-----RSEENIKYK--YGVTTELKIKYAMNKKRSRIKES 149
QY 376 L-----TNLYKKNKSGVDKNFNLDFKNNKNLDLDFFKNEKEYDDLCDCRY----- 423
DB 150 IYGAHNFPGNSYMEGDKGDKTG-----EEKDGEHKTSDKTDNGKANNLVMLDYETSSN 204
QY 424 -----TATIKSFLNGPAKN-----DVD-----TASQINVDNLGRF-----G 455
DB 205 GQPAATLNDLVLEFVTGHEGNSRKNSSNGNPGYDIDHKKTISSAI-IN--HAFLQNTVMKN 261
QY 456 CNYK-SNNEKSWNCTGFTNFKFCPTCEPPRRQTLCLGRTVLLHGHEDY-----KEH 507
DB 262 CNYKRRERDWC-----NWKDVCIPDRRYQJCMKELNVLNNTNTNHRDITPRKLY 316
QY 508 LLGASIIYEA-----QLLKYYKEKDNALCSIIQNSYADLADIIKGSDDIIKDYGGK 559
DB 317 LKRLIYDAAVEGDLLEKLNRY-NKD---FCKDIRWSLGDGFDIIMGTDMGIGYSKV 372
QY 560 MEENLKV-NKDKRNEESLKI FEKRWENKENVKNSAVLKNK-----ETCKDYDK 612
DB 373 VENNLRISIFGTDEKAQOR-----RKQWNESKAQIWTAMMYSVKRLKGNFIWICKLNV 427
QY 613 FQKIPQFLRFKFWGDDFCERKEKIKYSFESFKVEK-----KKKDCDENTCKKCEY 665

DB 428 VNIEPOLYRWIRWGRDYVSELPTEV---QKLKEKCDGKINYYTDKVKVPPCQACKSY 484
QY 666 KKWIDLKSEYKQVDKYTKDKNKKMYDNIDEVNKEANV-----YLKEKSECKOVNFD 720
DB 485 DQWITRKKQWDLVSNKFISVKNAB-----KVQTAGIVTPYDILKQELDEFNEVAPE 536
QY 721 DKIFNESPNEYEDMCKCKDEIKYLNIKYPKTKHDIYDITDFSDTFG-DGTPI----- 772
DB 537 NEI-NKRDGAYTEL-----VCSVEBAK-KNTQEVVTVNVDNAAKSOATNSNPISQPDSS 589
QY 773 -----SINANINEQSQKOTNTSGNSETSD-----SPVSHE-----PESDAINV 812
DB 590 KAEKVPDSTHGNV---SQDSSTTGKAVTGQNGNQTPAESDVQORSIDIAESVSANV 646
QY 813 E-----KLSGD-----ESSETRGILDINDPSPVTNNVNEVHSDASNTQGSV----- 852
DB 647 DPQKSVSKSDDTTASVTGIAEAGKNGLASRSPSESTVEANSPGDDTVNSASIPVVSGE 706
QY 853 -----SNTSDITNGHSESLNRTTNAQDIKIGRSG-----NEQSDNQENSHSS 896
DB 707 NPLVTPYNGLRHSKDNSSDGPASMANPDSNSK-----GETGKQDNDAKATKDSNSS 762
QY 897 DNSGSLT-----IGQVPSDNTQNTYDSQPHRDTFNALASLPSDDKINEIEG 944
DB 763 DGTSSATGTTDAVDREINKGVPEDRDKTVGSKDGGEDNSANKDAATVVGEDRIENSA 822
QY 945 FDSRD-----SENG-----RGDTTSN--THDVRRTNV 971
DB 823 GGTNDNRKNDTEKNGASTPDSKQSDATASLSTESLESSTESGDRITNDTNSLENKNG 882
QY 972 SERVNSHPFIRGMANNNAHGYIT-----QIENNGIIRGOESAGNSVNYKDNPK 1023
DB 883 KEKDLQKHPKSNDEPNSDQTTDAEGHDRDSIKND---KAERKHKMKDTFTKNTN 939
QY 1024 RSNFSEN---DHKNIOEVNSRDTKRVREIILKSKQKNCKNEYSMEYCTYSDERNSSP 1080
DB 940 SHLNNLNSLNGKLDIKYKRVKATREDIILMSVVKCNNNISLEYCNVEDKISS- 998
QY 1081 GPCSEERKELCQISDYCLKYFNYSIIBYNYNCIKSEIKSPYKCFKSEG-QSSIPYFAA 1139
DB 999 NTCSEKSKNLCCSISDFCLNYFDVYSYELSCMKKEFEDPSYKCTFKGFKIDKTYFAA 1058
QY 1140 GGLIVVIVILLSSASRMKSNSEYDIDGESNIETAFEE-NNYLNKLSRI-----FNQEOVE 1193
DB 1059 AGAL--LILLTASRMIKNDSE-----EATFNEFEYCDNIHRIPLMPNNIEHMQP 1108
QY 1194 TWISDYS 1200
DB 1109 STPLDYS 1115
RESULT 21
AAU76761
ID AAU76761 standard; protein; 972 AA.
XX AAU76761;
AC AAU76761;
XX 21-MAY-2002 (first entry)
DT 21-MAY-2002 (first entry)
XX Plasmodium falciparum erythrocyte binding protein EBP4 protein sequence.
DE Plasmodium falciparum erythrocyte binding protein EBP4 protein sequence.
KW Erythrocyte binding protein; EBP; malaria parasite; protozoacide;
KW vaccine; immune response inducer; Plasmodium falciparum merozoite; SABP;
KW salicylic acid binding protein; EBA-175.
OS Plasmodium falciparum.
XX WO200211756-A2.
PN 14-FEB-2002.
XX 07-AUG-2001; 2001WO-US024725.
XX

CC (RBC)) and a receptor peptide are claimed. AAR70103-25 are examples of
 CC these hybrid peptides. AAR70106 is a fusion of tumour necrosis factor
 CC receptor (in accordance with H Loetscher et al Cell, Vol. 61, 351-359)
 CC and Plasmodium vivax Duffy binding receptor. The use of cytokine receptors
 CC not normally found on RBCs means that the cytokine can bind harmlessly to
 CC the RBC without deleterious effect. The RBC protects the hybrid peptides
 CC from excretion from the kidney, and due to steric hindrance prevents the
 CC cytokines binding to a receptor in another cell. Glycophorin binding
 CC protein (GBP) 130 or GBP Homologue (H) are the pref. malaria parasite
 CC peptides used, others include EBA 175 (175 kDa erythrocyte binding
 CC antigen), Pwmsa (pre major merozoite surface antigen) and the Duffy
 CC binding receptor molecule (eg. exhibited by Plasmodium vivax). These
 CC peptides bind to pref. glycophorin A, B and C. sialo glycoproteins, found
 CC on the surface of RBCs. The hybrid peptides are thus used to lower the
 CC levels of free cytokines in the circulation to reduce pathological
 CC damage. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 16-OCT-
 CC 2003 to standardise OS field)
 XX
 SQ Sequence 1245 AA;

Query Match 11.5%; Score 743.5; DB 2; Length 1245;
 Best Local Similarity 23.1%; Pred. No. 2e-37;
 Matches 303; Conservative 192; Mismatches 466; Indels 353; Gaps 59;

QY 96 DIIPPSYSYRNDKFNLSNEDNSGNTNSNNTSISIGKONKQYTFIQKTHL---F 152
 DB 20 DSVCPQOKYIHPQNSISCTCKHKGTYLYNDPCPGQDTCRCEGSGSFTASENHLRHCL 79
 QY 153 ACGIKRKSIIKWI-----CREN-----SEKITVCPDRKIQLCVANFLNSR 192
 DB 80 SCSCRKEMQVEIISCTVDRTVCGCRKQYRHYWSENLPQCF---NCSLCLNGTVH-- 134
 QY 193 LETWEKPEKIFLISVNTAEKLLVNNKGDPSIFCNELRNSF----- 234
 DB 135 LSCQEKQNTV---CTCHAGFFLRENE---CVSCSNCKKSELECKLCLPOIENVKGTED 186
 QY 235 -----SDFRSSFIG-----DDMDPGGNTDRVKGYN-TKFSDDYKKNVKEKLN 277
 DB 187 SGTKDDFSITLINVHEGKKYLIILKRKLEKANNRDVCFNLFHPSQVNNVLLERTIETLLE 246
 QY 278 IKKEW--WEKNKANLNMHVIWNHKNISKECAIIPAEPOINLWIKENENFLMEKKRLF 335
 DB 247 CKNEYVKGNGKYLAKGHCV-----EEDNLERWLQGTNER----- 282
 QY 336 LNTKDKVENKKYKACFGGRLPCS-SYTFPMKKSQTQMEVL-----TNLYKKNSGV 387
 DB 283 -----RSEENIKYK--YGVTELKIKYAQMNGKRSRILKESIIYGAHNFNGSGYMEGKDG 335
 QY 388 DKNFLNDLPKNNKNDLDDFFKNEKEYDDLCDCRY-----TATIILKSLANGPAK 437
 DB 336 DKTG-----BEKGEHKTDSKTDNGKANNLMLDYETSSNGQPAGFLDNVLFVVTGHEG 390
 QY 438 N-----DVD-----IASQINVDLGRF-----GCNYK-SNNEKSNWCTGTFT 473
 DB 391 NSRKNSSNGGNPYDIDHKTKISSAI-IN-HAFLONTVMKNCVYKRRERDWDG----- 442
 QY 474 NKPGTCEPPRPTQLGRVYLLHRGHEDY-----KHLIGASIYEA-----QL 518
 DB 443 NTKKDVCIPDRRYQLCKMELTNLVNNTDTNPHRDIRTPKLYLKRKLIYDAAVEGDLILLKL 502
 QY 519 LKYKYEKENALCSIIQNSVADLADIIKGSDDIKDYVYKKMEENLNKV-NKDKRNEES 577
 DB 503 NTIRY-NKD---PCKDIRWSLGDPGDIIIMGTD-MEIGYSKVVENNLSIFGTDEKAQOR- 556
 QY 578 LKIPREKWMDEKNVWVMSAVLNK-----ETCKDYDKFQKIQFLWFKWEGWDDFC 631
 DB 557 -----RKQWNNESKQIWTAMWYSYKELKGNFIWICKLVAVNIEPQIYEWIRWEGRDYV 612
 QY 632 EKRKEKJYSPESFKVEK-----KKKDCDENTCKNCSYKKWIDLKXSEYKQVDKYT 684
 DB 613 SELPTEV---QKLEKCDGKINYTDKVKVPPQCNACKSYDQWITRKKQWDLVLSNKF 669
 QY 685 KDNKKWYDNIDEVKNKEANV-----YLKESKECKDVNFDDKIFNESPNEYEDMCKKCD 739

DB 670 SVKNAE-----KVQTAGIVTPYDILKQELDEFNEVAFENEI-NKRDGAYIELC----- 716
 QY 740 EIKYLEIKYPTKTHIDYDIDTFPSDTFG-DGTPI-----SINANINEQQS 783
 DB 717 -VCSVEEAK-KNTQEVVTVNDNAKSOATNSNPISQPVSSKAEKVPDSTHGNV---S 771
 QY 784 GKDTSVNGSETSD-----SPVSHE-----PESDAAINVE-----KLSGD----- 818
 DB 772 GQDSSTTGKAVTGDGQNGQTPAESDVQSRSDIAESVSANVDPQKSVSKRSDDTASVTGI 831
 QY 819 -ESSSETRGILDINDPSVNNVNEVDASNTQGSV-----SNTSDIT 859
 DB 832 AEAGKENLGASRPSSESTVEANSPGDDTVNSASIPVSGENPLVTPYNGLRHSKONSDS 891
 QY 860 NGHSESLARTTNAQDIKIGRSG-----NEQSDNQENSSHSSDSSGSLT----- 903
 DB 892 DGPAESMANPDSNSK-----GETGKGQDNMAKATKDSNSSDGTSATGDTTDAVDREIN 947
 QY 904 IGQVPSEDNTQNTYDSQNPHRDTPNALASLPDDKINEIEGFDSRD-----SENG--- 954
 DB 948 KGVPEDRDKTVSGKGGEDNSANKDAATVVGEDRIRENSAGGSTNDRSKNDTEKNGAST 1007
 QY 955 -----RGDTTSN--THDVERTNIVSERRVNSHDFIRNGMANN 990
 DB 1008 PDSKQSEDATALKTESLSTESGDRRTTNDTTNSLENKNGGKEDLQKHDFKSDNTPEE 1067
 QY 991 AHQYIT-----QIENNGIIRGOESAGSNVYKDNPKRNFSEN---DHKKNIOE 1039
 DB 1068 PMSDQTTDAEGHRRDSIKND---KAERRKHMNKDTFTKNTSHHLNNSNLSNGKLDIKE 1124
 QY 1040 YNSRDTKRVREBIIKLKSKONKCNNEYSMEYCTVYSDERNSSPGCSREERKKLCCQISDYC 1099
 DB 1125 YKTRDVKATREDIILMSSVRKCNNNISLEYCNSVEDKISS-NTCSREKSKNLCSSISDFC 1183
 QY 1100 LKYFENFYSIYNYCIKSEIKSPYKCFKSEG-QSSIPYFAAGILVIVILLSS 1152
 DB 1184 LNFYDVYSYELUSCMKKEFEDPSYKCTKGGFKIDKTYFAAAGALLILLILLIAS 1237
 RESULT 23
 AAR70231
 ID AAR70231 standard; protein; 1061 AA.
 XX
 AC AAR70231;
 DT 25-MAR-2003 (revised)
 DT 22-SEP-1995 (first entry)
 XX P. vivax DABP.
 DE DABP; duffy antigen binding protein; binding domain; merozoite; malaria;
 XX therapy; vaccine.
 XX Plasmodium vivax.
 XX
 Key Location/Qualifiers
 FH Domain 1..325
 FT /label= Binding_domain
 FT
 XX WO9507353-A2.
 XX
 PD 16-MAR-1995.
 XX
 XX 07-SEP-1994; 94WO-US010230.
 XX
 PR 10-SEP-1993; 93US-00119677.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Sim KL, Chitnis C, Miller LH, Peterson DS, Su X, Wellems TE;
 XX WPI; 1995-123427/16.
 DR


```

CC preventing malaria. (I) is useful as vaccine to prevent malaria or
CC infection by P. falciparum or P. vivax. The present sequence represents
CC the P. falciparum BBA-175 protein PfP2 domain sequence
XX
SQ Sequence 350 AA;
      Query Match      10.6%; Score 684; DB 5; Length 350;
      Best Local Similarity 37.2%; Pred. No. 1.9e-34;
      Matches 142; Conservative 70; Mismatches 124; Indels 46; Gaps 10;

Qy 446 INVNDLRGFGCNYKS--NNEKSNMCTGTFTNKPFGTCBPPTQTLCLGRVTLHHRGHEED 503
Db   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Qy 504 YKEHLLGASITYEAQLLKYYKKBKDENALCISIIONSYADLADIIRGSDIILKDYQKQMEEN 563
Db   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Qy 67  IKHILAIAIVESRLKRYKNKDDKEVKKIINFTFADIRDIIGGTDWNDLSNKLVGK 126
Db   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Qy 564 LNKVNDKKKNNEESLKIIFREKWDENKENVWKNVSAVLONKETCKVDVKFKQIPQPIRLWP 623
Db   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Qy 127 INTNSKTVVHRNKDKLFRDEWVKVIKKDVNVNVIISWVFKDKTVCKE--DIIENIPQFRWF 185
Db   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Qy 624 KEWGDDDFCEKREKEKIYFSESPKVECKKDCDENTCKNKCEYKWKWIDLKSEYEKQV--- 680
Db   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Qy 186 SEWGGDDYCDQTKMI---ETLKVCEKPCPCDDNCKSKNSYKEMIISKKKEEYNKQAKQY 242
Db   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Qy 681 DKYTKDKNKQMYDNIDEVNKEANVYLKEKSKCKDVNFDDKI FNEBSPNEYDMCKCKDE 740
Db   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Qy 243 QEVQKGNNYQY---SEFKSIKPVYLUKKYSEKSNLNFDEFKBEELHSDYFNKCTWCPE 299
Db   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Qy 741 IKYLNELIKPKTKHDIYDITFSDTFDGGDTPISINANINEQ--QSGKDTSTNGNSETSDS 798
Db   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Qy 300 VK-----DVPISIRN-NEQTSQEAPEENVEIAHRTET 332
Db   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:

Qy 799 P-VSHEPESDAAINVEKLSGDE 819
Db   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Qy 333 PSISEGPKG---NEOKERDDDD 350
Db   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:

```

RESULT 26	
AAR13456	
ID	AAR13456 standard; protein; 778 AA.
XX	
AC	AAR13456;
XX	
25-MAR-2003	(revised)
DT	17-DEC-2001 (revised)
DT	07-NOV-1991 (first entry)
XX	
DE	Duffy receptor.
XX	
KW	Malaria; parasite; vaccine.
XX	
OS	Plasmodium knowlesi.
XX	
XX	
Key	Location/Qualifiers
FT	1. .713
FT	/label= exon 1
FT	265. .286
FT	/label= peptide 3
FT	525. .569
FT	/label= repeat region
FT	/note= "nine repeats"
FT	525. .529
FT	/label= repeat unit
FT	/note= "SSD (Q/H)T"
FT	738. .739
FT	/label= exon 2
FT	763. .764
FT	/label= exon 3
FT	765. .778
FT	/label= exon 4
FT	765
FT	Peptide

FT	XX	/label= C-terminal peptide
PN	XX	USN7554837-N.
XX	XX	
PD	XX	23-JUL-1991.
XX	XX	
PF	XX	20-JUL-1990; 90US-00554567.
XX	XX	
PR	XX	20-JUL-1990; 90US-00554837.
XX	XX	
PA	XX	(USSH) NAT INST OF HEALTH.
PA	XX	(USDC) US SEC OF COMMERCE.
XX	XX	
PI	XX	Miller L, Adams J, Kaslow D, Xiangdong F;
XX	XX	
DR	XX	WPI; 1991-260184/35.
XX	XX	N-PSDB; AAQ13316.
DR	XX	
XX	XX	Plasmodium Duffy receptor proteins - used in vaccines against malaria and
PT	XX	for producing antibodies which prevent malaria infection.
XX	XX	
PS	XX	Disclosure; Fig 1; 63pp; English.
XX	XX	
CC	CC	The sequence was deduced from clones isolated from a P. knowlesi genomic
CC	CC	library. The sequence covers 4 exons and has a 22 amino acid
CC	CC	transmembrane segment (exact posn. not given) followed by 45 AAs at the C
CC	CC	-terminus. Nine repeats of the pentapeptide SSD(Q/H)T occur 5' to the
CC	CC	transmembrane segment. Two regions of high Cys content are separated by a
CC	CC	Pro-rich region. The C-terminal peptide and peptide 3 were conjugated to
CC	CC	KLH and used as immuno- gens. The gene can be used to express recombinant
CC	CC	receptor for use in the prepn. of vaccines against malaria. See also
CC	CC	AA13457. (Note: Revised entry submitted to correct the patent number
CC	CC	format of US Government-owned NTIS applications to prevent clashes with
CC	CC	ongoing US granted patent numbers. For further information please visit
CC	CC	the Derwent web site at www.derwent.com/dwpi/updates/ntis.us.html .)
CC	CC	(Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to
CC	CC	correct PA field.)

[illegible]

```
Db 393 DGEKADPKDIEV--KGQDTPDORSQSLGPHGHTDERATLGETHMEKDTETAGGSTL----- 446
Qy 926 TPNALASLPD-----DKINE-----IEGFSSRDSSENGRGDTTISNTHDVRRTIVSE 973
Db 447 TPEQNVSVASDNGNVPGSGKQNEGATALSABSLKSNESVHKTIIDNTHTHGLENKGGNE 506
Qy 974 RRVNSHDFIRGMANNA----- 991
Db 507 KDFQKHDFMNDMLNDQASDHTSSDQTSDDHTSSDQTSDDHTSSDQTSDDHTSSDQTS 566
Qy 992 -----HHOYITQIENNGIIRGOESAGNSVNYKDNPKRSNPFSSNDHKKNIQEVNSR 1043
Db 567 DQITDTEGHR--DNVRNPEIKSEDSKSGDFMRNSNELYSHNNLNRRKLARDQYEHR 624
Qy 1044 DTKRVREIIKLSKQNKNNMEYSMEYCTSDERNSSPGCSREERKKLCCQIISDYCLKYP 1103
Db 625 DVKATREKIILMSVKNKNNRASVKYCNVTIEDRMLS-STCSRRRKNLCCSISDFCLNYP 683
Qy 1104 NFYSIEYNCIKSEIKSPYKCFKSEGOSSIPYPAAGGILVIVILLSSASRMKSNEEY 1163
Db 684 ELXSYEFYNCKBEFEDPSYECFTKGSSTGIVVPATGGAFLIILLFPASWNAASNDYEE- 742
Qy 1164 DIGESNIEATFEE 1176
Db 743 -----EATFDE 748

RESULT 27
ID ADO69969 standard; protein; 2459 AA.
AC ADO69969;
XX
XX
DT 29-JUL-2004 (first entry)
XX
DE Plasmodium falciparum virulence factor var O protein.
XX
XX antimalarial; vaccine; gene therapy; virulence factor; var O; Plasmodium;
KW malaria.
XX
XX Plasmodium falciparum.
OS
XX WO2004037856-A2.
XX
XX 06-MAY-2004.
XX
XX 24-OCT-2003; 2003WO-EP013341.
XX
XX 25-OCT-2002; 2002CA-02409897.
XX
XX (INSP ) INST PASTEUR.
XX
XX Puijalon O, Le Scanf C, Lavergne A, Bentley G, Badaut C;
PI Igonet S;
XX
XX WPI; 2004-399926/37.
XX
XX N-PSDB; ADO69968.
XX
XX New virulence factor var O polynucleotide of plasmodium falciparum,
PT useful in preparing a composition for treating or preventing malaria.
XX
XX Claim 9; SEQ ID NO 2; 84pp; English.
XX
XX The invention relates to an isolated or purified virulence factor var O
CC polynucleotide comprising a nucleic acid sequence that is 65, 80 or 95%
CC identical to a sequence comprising 7378 base pairs (bp) or its fragment.
CC The polynucleotide is useful in preparing a composition for treating or
CC preventing Plasmodium species related disease, particularly malaria. This
CC sequence corresponds to the virulence factor varO protein.
XX
XX Sequence 2459 AA;
```

```
Query Match 10.0%; Score 651; DB 8; Length 2459;
Best Local Similarity 23.9%; Pred. No. 3.3e-31;
Matches 245; Conservative 120; Mismatches 307; Indels 352; Gaps 43;

Qy 75 ENKSVKKKRSLSF-----INNKTSDYDIIPPSYSYRNDKFNSLSENEDE-----SGNTN 123
Db 1550 ENSSTQKDLPEAFDPCPKIIGDRNCPCPKPEKYCVKTAHDIRKENGKGNFDNLKGNVN 1609
Qy 124 SNN-----FANTSISIGKONKQYTFIQKRTHLFACGIKRKS I-- 161
Db 1610 TYNDNCKNAKREDYANQNGETCKFEVSWSSIGIINNENE-----STGRDFKIGE 1660
Qy 162 KWICRENSE--KITVCVPDRKIQLCVANFNSRLEWKEFEKFLISVNTAEKLLYN--- 216
Db 1661 VMECKNETTGGKNKVCVPPRRKDMC-----LKKLODIRVDDISDSSTLLKBIQE 1709
Qy 217 --KNEGKD-----PSIFCNE-----LRNSFSDRSPSFIGDD-----MDFGGNTDR 254
Db 1710 VAKNEGNDIIRNLPKYPCNEDVICYMKYSFADLGDIVRGTDKYKDVIGSISGNNSEQ 1769
Qy 255 VKGYINTKFSDYK-----EKNVEKLNNIKGEWERNKANLNMHIMVN-----HKGNISKE 305
Db 1770 IBEENLKTIFENIQKTDENFQKYTNLELFRSAWWDANRDKIWKAMTCNAPDEAKIYITKE 1829
Qy 306 CAII-----PAEE-----POINLMIKEWNEFLMEKKRLFLNIKDKC----- 342
Db 1830 GGYISPITWTQNHCHGNDPPDYIYIPLRWISEWSYCYLAQKD-FLETMKNCENCK 1888
Qy 343 -----VENKYEACFGGRLPCSSYTSFMKSKTOMEVLTNLYK-----KONS 385
Db 1889 KNDNTCEBQTKYGAC-RDCKKCEEYKFDKWAQAFETQNKAYKAYIKNATTSRGRHSN 1947
Qy 386 GVDKN--NFLNDLFKKNNKNDLDD-----PPKNEKEYD 416
Db 1948 GIDEDIKKFVEKLEQNCQKNSVDTDADKYLEGGSVCRFRKFKVTDTHKKNVAFHTPLSYK 2007
Qy 417 DLCDCRYTATIKSFLNGPAKN-DVDIASQINVDLRGFG-----CNYK 459
Db 2008 EHCEC-----AKNFDPLDECPVDNNECKYTGIGSCPEKNFHKLEERTNYV 2053
Qy 460 SNNEKSNCTGTFTNKFPGTCPPROTCL-GRTYLLHR-GHEEDYKCHLLGASIVEAQ 517
Db 2054 LNNK-----SNKNSAIVPPRRRQQLQNLNRLNLRNKEKSKFEGILLISAASEAK 2104
Qy 518 LLKYKYKEKDNALCSIIQNSYADLADIIKGSIIKDYTGKKMEENLNK-VNKRKRNEE 576
Db 2105 MLTEQYRENPAAKAL-QAIKYSPADIGNIIGDDII-----GNVISVOLKLNKLNKKNITS 2159
Qy 577 SLKIFREKWDENKENVWQMSAV-----LKNKETCKDYKFKQIPOPFLRWFKEWDDPCE 632
Db 2160 TL-----WWEANKEKINWAMMCLHWRKKTKATSPSHDNIDKEDQFLRWFOEWGENFCA 2213
Qy 633 KRKEKIYSPESFKVECKKDC-----DENTCKNKSEYKWKWDLKSEYQVQDKYTK 685
Db 2214 TRKE-----LYEKLNNCKSVCEKNASKGNVNLKCTKACEYKSVLKKTYEYIQLQKYDK 2270
Qy 686 DKKNKMYNDIDEVKNKEANVYLKESKECKDNVFDKIFNESP-NREYED-----MCKK 737
Db 2271 EFNKTL-----NNKNALEFL--NVQCISEYFSDSKWESPYDTFDDTLKGTVDCKK 2320
Qy 738 CD----- 744
Db 2321 HEPTPAIKPSKPASPEDKKLVDSPLIPIQOPQSNNTSDILATTIPFGIALGSAFL 2380
Qy 745 NETKYPKTHDIYDI-DTFSDTFGDGTPIISINANI--NEOQSGKDTNTGNETSDSPV 800
Db 2381 FLKKKPKSPVLLRLVLDVHKDYGTPPKSSNRYIYPASDRHKGK----- 2425
Qy 801 SHEPESDAAINVEKLSGDESSSETRGILDINDPDSVTNNVNEVHDASNTQGSVNTSDITN 860
Db 2426 -----TYIYMEGDSGDEKYAFMSV-----TTDVT 2451
Qy 861 GHSE 864
```

```

Db      2452 SESE 2455

RESULT 28
ADR43499
ID      ADR43499 standard; protein; 3056 AA.
XX
XX
AC      ADR43499;
XX
XX
DT      04-NOV-2004 (first entry)
XX
XX
DE      CSA-binding malarial variant surface antigen.
XX
KW      antimalarial; gene therapy; vaccine; Plasmodium falciparum; malaria;
KW      variant surface antigen; glycosaminoglycan; chondroitin sulphate A;
KW      immune response; erythrocyte; placental parasite; prophylaxis;
KW      pregnancy-associated malaria; diagnosis.
XX
XX
OS      Plasmodium falciparum.
XX
XX
PN      WO2004067559-A1.
XX
XX
PD      12-AUG-2004.
XX
XX
PF      30-DEC-2003; 2003WO-DK000938.
XX
XX
PR      27-JAN-2003; 2003DK-00000102.
XX
XX
PA      (UYKO-) UNIV KOBNHANS.
XX
XX
PI      Theander TG, Salanti A, Hviid L, Staalso T, Jensen ATR;
PI      Lavatsen T, Dahlbaeck M;
XX
XX
XX      WPI; 2004-594166/57.
DR      N-PSDB; ADR43498.
XX
XX
New polypeptides and encoding nucleic acid molecules useful for
PT      diagnosing, preventing or treating pregnancy-associated malaria or for
PT      identifying agents capable of preventing or treating pregnancy-associated
PT      malaria.
XX
XX
PS      Claim 1; SEQ ID NO 2; 138pp; English.
XX
XX
This sequence corresponds to a Plasmodium falciparum (malaria) variant
CC      surface antigen which binds to the glycosaminoglycan chondroitin sulphate
CC      A. The protein is designated vsa2csa. The protein is capable of inducing
CC      an immune response against a molecule expressed on the surface of an
CC      intact erythrocyte infected by a placental parasite. Sub-sequences
CC      comprise at least one B-cell epitope or one or more CAG-binding motifs
CC      and do not comprise a CIDR domain or DBL-gamma domain. The protein is
CC      gender specifically recognized or is recognized in a parity dependent
CC      manner. The polypeptide or nucleotide sequence is useful for
CC      manufacturing a composition that prophylactically or therapeutically
CC      reduces the incidence, prevalence or severity of pregnancy-associated
CC      malaria in a female subject. The composition and methods may also be used
CC      for diagnosing the above condition or for identifying agents capable of
CC      preventing or treating pregnancy-associated malaria.
XX
XX
Sequence 3056 AA;
Query Match          9.8%; Score 637.5; DB 8; Length 3056;
Best Local Similarity 22.9%; Pred. No. 3.1e-30;
Matches 297; Conservative 165; Mismatches 462; Indels 371; Gaps 66;

Qy      1 MKGYFN-IYFLIFLYNIRINESIIGRTLYNRQDESSDISRVNSP----- 47
Db      1854 MSYNYNKIY-----RKSNESEGDGYSMIMEFTVIDLNKRCNG 1893
Qy      48 ELNNNH-----KTNLYSDYEDVNNKLNINSFVKNKSVKKRSLSPNNKTKSYDIIPPSY 102
Db      1894 BINGNYICCSCKNIGENSTSGTVNKLQKK--ETQCEDNKGPLDLMNKLNMVD---PKY 1948

103 SYRNDKFN--LSENEDSNGTNSNNTFANTSEISIG-KDNKOYTF---IQKRTHLFAC-- 154
1949 SEHKMKCTEYVLEHVEE-----OLKEIDNAIKOYKLYPLDRCFDDKSKMKVCDL 1997
155 -----GKRKS-----IKW-----ICRESEKIT-VCVPRDKIOLCVANFLN--SRLETMEK 198
1998 IGDAIGCKKHKTCLDELDEWNDVMDRPYNKYKGLIPRRRQLCFSTRVGRPANLNJKE 2057
199 FKEIFLISVNTKALYN-KNEGKOPSIFCNELNSESDFRSPFIQDDMDFGGNTDRVKG 257
2058 FKEILKGAQSEGKPLGNYNEDKDKKALAMKNFSYDYIYIKGSDM----- 2106
258 YINTKFSDDYK--EKATVEK-LNNIKK--EMWEKKNKALNWNMIVNHK--GN--ISKCAI 308
2107 LTNIQDKIKRKLRLLEKETNNTKEVDVDDWETNKKSIWNAMLCGYKSGNKKIIDPSWCT 2166
309 IPABE--POINLWIKWENENFLMKRFLNLIKOC--VEN-KKYEACFGGRLUPCSSYT 363
2167 IPTTETPPQFLRWIKWGTNVCIQEBEHKEVYKSCSNVTNLGAQESSEKNCSTSEIKKYQ 2226
364 SFMEKSKTOMEVLNLYKKNSGVD--KNFLNDLFPKKNKNDLDDPFKNEKEYDDLLCDC 421
2227 BWSRRSIQWEAISGEY--KKYGMDEFKNTPKNIKEPDANEPNANEYL---KKHCSKCP 2282
422 -----RYTATIIKSFNGPAKNDVDIASQIN-----VNDLRGF 454
2283 GFNDMQEITKYTNIGNEAFKQ--IKEQVDIPAELEDVIYRLKHHEYDKGNDYICNKYKNI 2340
455 GCNYKSNNEKSWN---CTGFTFNKPGTCEPRPRQTLCLGRTYLLHGRHESD----- 503
2341 NVNMKGNDDTWTDLVKNSSDINK--GVLLPFRKRLFL-----KIDESDICKYKRD 2391
504 --YKEHLLGASIVYEAQLLYKYEKENALCSIIQNSVADLADIIKGSDIK---DYVG 557
2392 KLFKDFIYSSAISVEERLKKVYGEA-KTKVVMHAKYFADIGSIKGDMMENNSSDKIG 2450
558 KMEENLNKVNKKRNEESLKIIFREKWDNENKENVKVMKSAVLNKK--ETCKDYDFKQK 615
2451 KILGDGVG-----QNEK-----RKWWDNKKYHIESMLCGYKHAGVNTSENDKMLD 2498
616 IP-----QFLRWFKWGGDDPCFKRKGKIYSFESFKVECKKQC-----DENTCKNKC 662
2499 IPNDDDEHQFLRWFPQEWTFENFCTKRNE---LYENNVTACNSAKCNTSNGSVDKKECTEAC 2555
663 SEYKWKDLKSEYKQVDKYTKDKNKKMYNDIENVKNKEANVYLKESK---ECKOVNF 719
2556 KNSYNFILIKKKYQSLNSQY--DMNYK-----TKAEKKSPPYFKDKCNGSECSLSEYF 2609
720 DDKIFNESPNEYED-----MCK-----KCKEIKYLNKIK 748
2610 XDTRFNKNPYETLDDTEVKNNCKMCKPPPPASNNNTSDILOKTIIPGIALALGSIAPFKMK 2669
749 YPKTKHDIYDI-DYFSDTFGDTGPISINANI---NEQOSGK-----DTSNTGNS----- 793
2670 KPKTPVLLRLVDLPKDGIGIPTKSSNRYIPYASDRYKGYKTYIYMEGDTSGDDDKYIWD 2729
794 -ETSDSPVSHPEPDAIN-----VEKLSGD 818
2730 LSSSDITSSSESEYEEVDINDIYVPSFKPKYTFIBLVLEPSKRDFTNTSSGDTFTNKLTTDD 2789
819 ESSSETRGILDINDPSTVNNVNE---VHDASNTQGSVNTSD--ITNGHSE----- 864
2790 EWNQKQKDFIE---QYLNQIKQDFILHDSWDEKFFITQDRFLDSHSEVYTNIDWNV 2845
865 -SSLNRTTNAQD-----IKIGRSGNEQSDNOENSSHSDNSGSIITIGQVP 908
2846 PENINRITNNMDPKYCSNNMYTGTDLINDSLANGQYDIYDEMLKRENE---LFGTYH 2902
909 SEDNTQNTYDSNP-----HRTFPNALASLPDSD---KLINEIEGPDSS 948
2903 TKYTFNFSVKQTPSDPIINQLDLHYKWDIKHRDICEQWKT--KEDMLYKLNWNNMERK 2960
949 RDSENGRGDITTSNTHDV--RRTNIVSRRRVNSH-----DFI-----RNG 985

```

```

Db      2961 EYLLDIQPSLDDIHKINDEYNIISTNNIYDHPQSQETPLQLLGSTNIISFYITTEQNG 3020
Qy      986 MANNNAHQYITQENNGIRGQ---BESAGNSVN 1017
Db      3021 LRTNISMDTYIDETNNNNVATSIIGDDQMSNVN 3055

RESULT 29
AA62142
AC      AAB62142 standard; protein; 3542 AA.
AC      AAB62142;
XX
XX      29-MAY-2001 (first entry)
XX
XX      P. falciparum FC3.varCSA protein.
XX
XX      FC3.varCSA protein; chondroitin sulfate A; CSA; var gene; PFEMP1;
KW      erythrocyte membrane protein 1; parasitized red blood cell; PRBC;
KW      malaria; protozoacide.
XX
XX      Plasmodium falciparum.
XX
XX      W0200116326-A2.
XX
XX      08-MAR-2001.
XX
XX      01-SEP-2000; 2000WO-US024195.
XX
XX      01-SEP-1999; 99US-0152023P.
XX
XX      (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX      Scherf A, Miller LH, Gamain B, Baruch DI, Buffet P, Scheidig C;
PI      Gysin J, Pouvelle B, Fujii N, Smith J;
XX
XX      WPI; 2001-235109/24.
XX
XX      N-PSDB; AAF57301.
XX
XX      Novel FC3.varCSA protein, useful for modulating parasitized red blood
XX      cell binding, sequestration and onset of maternal malaria.
XX
XX      Claim 12; Page 63-71; 78pp; English.
XX
XX      The invention relates to a P. falciparum FC3.varCSA protein, that is
XX      capable of binding to chondroitin sulfate A (CSA). The var gene and the
XX      corresponding P.falciparum erythrocyte membrane protein 1 (PFEMP1)
XX      modulate adhesion of parasitized red blood cell (PRBC) to CSA. The
XX      protein and the encoding gene are useful for treating and preventing
XX      maternal malaria in a patient identified at a risk for contracting
XX      maternal malaria or in a patient afflicted with maternal malaria. The
XX      present sequence represents the P. falciparum FC3.varCSA protein
XX
XX      Sequence 3542 AA;

Query Match          9.3%; Score 601.5; DB 4; Length 3542;
Best Local Similarity 20.8%; Pred. No. 6.9e-28;
Matches 303; Conservative 192; Mismatches 466; Indels 499; Gaps 68;

Qy      56 NIYDSYEDVNNK---LINSFVENKSVK---KRSLSFI-----NNKTKSYD- 96
Db      1118 NIISDKYKELHEQAMSVNSGIGBASSTAKNHIDRNVEIFSELYQQGKSNKSGTSDE 1177
Qy      97 --IIPPSYSYRNDKFNLSNEDNSGN---TNSNNFANTSEISIGKDNKYTFIQK-RT 149
Db      1178 SAVIGTWTYEN-----VGAYLHDTGTFDDCQSQNEFCD--EKSDGDKNEKYAFRDKPD 1230
Qy      150 HLFACGKIKR-----KSTKWKICRENS----- 169
Db      1231 HDGACGCKSGKPTRVQIKTKKAAEBKDTECKTVNDILKENDGKKQVEDCHPKKNSNGYP 1290
Qy      170 -----EKITVCVPDRKIQLCVANFLN-----SRLETMEKFEFLISVNTAE--KL 213

```

```

Db      1291 DMOCGNINLVEDPRVCPMPRRQKLCVHFELANDNEIKKLSQVNLKEAFISAAAEFFSW 1350
Qy      214 LYNKN-----EGKOPSIFCNELRNSFSDFRSSFISGDDMDFG-GNTDRVKGYI 259
Db      1351 YYYKSKDGBEGNLDKELKEGKIPPAFLRSNFFYFGDYRDLFTGDTISKGHGEKSLKEQI 1410
Qy      260 NTKFSDYYKEKNVEKLN---NIKKEWKNKANLWNHMIVNHKGNISKECAIIP----- 310
Db      1411 DSLF-----KNGDQKSPNGKTRQEWWTSHSHEIWAML-----CALVKIGAKKD 1454
Qy      311 -----AEEPJNLWIKENENFLMEKKRLFLNLKDKCVENKK 347
Db      1455 DFTENGYNNVFPKSDKSTLEBPAKRPQFLRWITWYDYCYTRQKYLKDVQKCKSNDQ 1514
Qy      348 YEACFGGCRLPCCSSYTSFMKSKSTQMEVLTNLYKKNSGVDKNNFLNDLFFKNNKNDLDD 407
Db      1515 LK-CDTECNKKCEDYVYMKKKKEWIP-QDKYKDER---DKGRF----- 1554
Qy      408 FFKNEKEYDDLCDCRYTATIISFLN-----GPAKNDVDIASQINVNDL----- 451
Db      1555 ---DRQHIGVMVTDYTGNTATDYLNRKFTASCGDKPGSASV-QRNIQLLEKQAYYDAD 1609
Qy      452 RGFGC-NYKSNNKESWN-----CTG-----TFTNKFPGTCE-----PP 483
Db      1610 KHGCTKFIENDDKYTNISSKOKCKGLVKEANTGAIKWQNKGNPNYNNLKELTEDVLFPS 1669
Qy      484 RRQTLCLGRTYLLHRGH-----BEDYKEHLGASIEAQLLYKYKEKEDNAL----- 531
Db      1670 RRLRICF---HALDGNVTDPEVKDENGRLKMEVAATEGYNLGOYYKEKEKIKTSD 1726
Qy      532 -----CSIIQNSYADLADIIGSDIHKDYGGKMEENLNKV-NKD-----KRNNEE 576
Db      1727 AHKYSYEVPPCSAMKYSFYDLRDIILGIDNLED-ERQKTEENLKIFNKNGTSGVGKSDS 1785
Qy      577 SL---KIREKWNDEKENVKMSAVLK-----NKETCKDYDFQKIP----- 617
Db      1786 TTGNPGSTARKFPWNNKBCVWNAMICGYKRGDRDNGNSARSBDLKKCGSVSDDDY 1845
Qy      618 -----QFLRWFKWGGDDFCBKEKEKIYSPSPKVECKKDCDEN-----TCNK 660
Db      1846 PMGKNRDEGTAYQFLRWFAEWGEDFC-KHKEK--ELEKLVGACNDYTCGNEDKRXKCTD 1902
Qy      661 KCSEYKWIIDLKSEYEKQVDKYTKDNKKMYDNIDEVRNKE--ANVYLKEKSKCKDVNF 719
Db      1903 ACTQYKKFISEWKPOVEKQIKKYGENK-KIYSEHPVAKDAEDAREYLDKQLK----- 1954
Qy      720 DDKIFNESPNEYEDMCKCDEIKYLNKIKYPKTKHDIYDIDTFDTFGDCTPISINANIN 779
Db      1955 --KICENKSGDCBYKCMK-----DVSTQRITDGNSSQNMPSASLD 1990
Qy      780 EQOSGKDTNTGNSETSDSP--VSHEPESDAAINVEKLSGDESSESTRTGILDINDP-SVT 836
Db      1991 DEP--KEVEGKNCQVPRGPPVRVRRTPSPRVSLSISKATA--SKKEAKTAPTKQPKKVE 2046
Qy      837 NNVNEVHDASNTOGSVSNTSDITNGHSSESLNRTTNAQDIKIGRSNGEQSDNQNSHSS 896
Db      2047 NLTTENRAQTRTRAAQQRKRTS-----TATTTESD--VGTWKAILLSNKPDSRGGI 2097
Qy      897 DNSGSLTIGQVP-----SEDN-----TONTYDS----- 919
Db      2098 EGCNPKTYGQYPKWGCIVGKSENENGI CWPPRKKLCINNIQYLNAYETENKKNNDI KEA 2157
Qy      920 -----QNPHRDTFNALASLPD-----DKINEIEGFSDSRDS 951
Db      2158 FIKCAAIETQFLWLVIIENPAEENELNGTI PDEFKRIWYTYGYDKOMFGFGTIDISNDK 2217
Qy      952 ENGRGDTTNTDVRNTIVSERRVNSHD-----FIRNGMANNNAHQ 994
Db      2218 K-----IITVTSV--TTILNENKCKDKKDEELRKIFWEKNKRFIWEIGMYGLTYH- 2269
Qy      995 YITQIENNGIIROQESAGNSVNYKD---NPKRSNPFSSENDHKNIQENSRDTRKVRRE 1050

```

Db 2270 -LTD-EN-----EKEKIRDNYQVNDMTKLTPLSLEEFVKRPQFLRWFTWEAEFCNKKE 2321
 Qy 1051 BIKLS---KQKCNNEYSMEYCTYSDBRNSPGPCSRERKCLCCQISDYCLKYENFYS 1107
 Db 2322 QLLKLEAGCKEYECNG-----SNDGKTQEC---AEACTYQNF-- 2356
 Qy 1108 IBYNCIKSEIKSPRYKCFKSEGGSSIPYFAGGILVIVILLSSASRMGKSNBEYDIGE 1167
 Db 2357 -----IKKWTEYERQREKF-----KKDKGKKYKIDYPSTE 2387
 Qy 1168 SNIATPBENNYLN-KLSRI 1186
 Db 2388 RDIEKATCAHEYLNNKLKEL 2407
 RESULT 30
 AEB22180
 ID AEB22180 standard; protein; 3147 AA.
 XX
 AC AEB22180;
 XX
 DT 22-SEP-2005 (first entry)
 XX
 DE Codon optimized P. falciparum VAR4.
 XX
 KW antimalarial; vaccine; protein purification; pharmaceutical; diagnostic;
 KW plasmodium falciparum infection; infection; VAR4.
 XX
 OS Plasmodium falciparum.
 XX
 FN WO2005063804-A1.
 XX
 PD 14-JUL-2005.
 XX
 PF 28-DEC-2004; 2004WO-DK000922.
 XX
 PR 30-DEC-2003; 2003DK-00001954.
 XX
 PA (UYKO-) UNIV KOBENHAVNS PANUM.
 XX
 PI Jensen ATR, Hviid L, Jorgensen L, Lavsten T, Magistrado P;
 PI Nielsen MA, Salanti A, Staalso T, Theander TG;
 XX
 DR WPI; 2005-522559/53.
 DR N-PSDB; AEB22179.
 XX
 PT New isolated VAR4, VAR5, and VAR6 polypeptides isolated from Plasmodium
 PT falciparum, useful as vaccines for treating or preventing severe malaria.
 XX
 PS Disclosure; SEQ ID NO 8; 193pp; English.
 XX
 CC The invention describes an isolated polypeptide comprising an amino acid
 CC sequence selected from at least one of SEQ ID NOS 2, 4, and 6 (3553, 2994
 CC and 3346 amino acids) for use as a medicament. The polypeptides, nucleic
 CC acids and vectors are useful as vaccines for preventing and/or treating
 CC malaria. The malaria is preferably severe malaria caused by Plasmodium
 CC falciparum. An in vitro diagnostic method is useful for detecting anti-
 CC malarial antibodies. The polypeptides are useful for generating a vaccine
 CC against severe malaria, where the vaccine comprises antibodies generated
 CC against the polypeptide which have cross-reactivity to parasites causing
 CC severe malaria and the ability to inhibit adhesion to endothelial cells.
 CC The polypeptides are also useful in screening for inhibitor molecules
 CC capable of inhibiting binding of any of the polypeptides to a receptor
 CC expressed on endothelial cells. This is the amino acid sequence of
 CC Plasmodium falciparum Var4 that is transcriptionally upregulated in
 CC Plasmodium following increased antibody recognition and is encoded by a
 CC codon-optimized polynucleotide.
 XX
 SQ Sequence 3147 AA;

Query Match 9.1%; Score 592; DB 9; Length 3147;
 Best Local Similarity 20.0%; Pred. NO. 2.3e-27;
 Matches 315; Conservative 205; Mismatches 511; Indels 544; Gaps 68;

Qy 39 SDISRVNPELNNHKTNIYDSYEDVNNKLINSFVENKSVKKRSLSFNNKTKSYDII 98
 Db 1252 TDVAKILQGEAN-----ETWLNSSNGNDKDSKLGKAEGBD-- 1289
 Qy 99 PPSYSYRNDKFNSEHEDNSGNTNSNFANTSISIGKONKYQYTFQKRTHLFACGIKR 158
 Db 1290 ---YS-RGGTSPDFNNLCLGITHKSHANDSQQPCYCKQQRNFV----- 1331
 Qy 159 KSIKWI CRESEKIT--VCVPDRKIOLCVANFLNSLETM--EKFEKIF----- 203
 Db 1332 -GTEWSPFADNHRKTHPEAYMPPREHICTSN-----LEYLIHRKKPIIEGDPNKIHS 1385
 Qy 204 ----LISVNTA---KLLYNKNEG-KDPSIFCNELRNSFSDFRSSFIGDDMDGFGNTDR 254
 Db 1386 LLGDVLLAAKYEAEINIKLYEENNRKQDEGICRAMKYSFADIGDIIRGKDM-WIENDA 1444
 Qy 255 VKGIYN-----TKFSD-----YYEKQVVEKLNNIKKEWKEKANLWNHMIWNHG-NIS 303
 Db 1445 KRLQTNLKEIPTKIKERTGGTTYNEDNDPVL-KLRADWWEANRAKVMKAMCKTNGVDIT 1503
 Qy 304 KECALIPAEB--POINLIWIKENENFLMEKKRLFLNIKDCVS---NKKYEACF----- 352
 Db 1504 CDSHTPLDDYIPORLRWMTAEAWCYKASQETKLEEKCSQCKSGKGGNECTREYKE 1563
 Qy 353 -GGCRLPCSSYTSFMKSKSKTQMEVLTN---LYKKKNSGV-----DKN--NFLN 394
 Db 1564 CNDCKQACEEYKRIKTWADQWKVISNKYEDLYKKAQNPNAVLKDNKDKKVDPLT 1623
 Qy 395 DLFKKNN-----KNDLDD-----FFKNEKE-Y 415
 Db 1624 QLOKANNGEKTGVHTVYSTAAGYIHOEARTRECOEREFCDKNGIDNTSYAFKDPHG 1683
 Qy 416 DDLCDCRYTATIKSFLNGPAKNDVDITASQINV-----NDLRGFG---CNVK-----S 460
 Db 1684 ATACDC-----INRSQTEEPKKEENVESACKIVVEVLSRPRDKTTGTDHCHNKPYPRK 1738
 Qy 461 NNEKSNWCT-GTFTNKPFGCEPPRRQTLG-LGRTYLLHRGHEEDYKHEHLGASIEAQL 518
 Db 1739 ENYFGWNTCPQFSKHAGACMPERRIKLCVINQYLNKESPEELKAFQCAAIETW 1798
 Qy 519 LKYKYKEKDNALCISIIONS-----YADLADIIGSDIHKDYGK---KM 560
 Db 1799 LMQYKKDKNGVAQAALNSGTIPDDFKRQMYFTFGDYRDLCLDTDISSKADTSTGVGKV 1858
 Qy 561 EENLKV-NDKKNEESLKFREKWDENKENVKWSAVLKNKEICK--DYDKFQKI- 616
 Db 1859 KINIDSVFQKIDITNVEQ-----RKPWNGKNAEATWDGMLCALSYNTTKNNMDYNAHTKLN 1914
 Qy 617 -----POFLRPFKEWGDGDFCEKREKIYSFESFVKECK----- 649
 Db 1915 PTYGYNAIKSELEDFVNRPOFLRWFTWSEDFCTERSIKIKELCTKNDCTVSSGTSDA 1974
 Qy 650 --KKDCDE---NFKCNKSEYKQKWDILKSEYKQVDKYTKDNKKMYDNIDEV-KNKE 702
 Db 1975 TGNKTCDDKDCBCKRACCTYKTLKNWKYTKYQSKKYFDDKRKELYSIDDVASSTQ 2034
 Qy 703 ANVYLKESK-----ECKDVNFDKIFNESP-----NEYDMCKEC----- 738
 Db 2035 AYQYLHAQLKKLKGADCKMCGESKETQGPNSHSHMPASLDDEPEVNGKCNCKVK 2094
 Qy 739 -----DEIKYLN-----EIKYPKTKHD 755
 Db 2095 HRPOPPLALPPAPSGPAAEQIEHNRGSRERDQGLPARPPPPQAAPPQPKPRT 2154
 Qy 756 IYDIDTFSDTFGDG-----TPISINANINBQQSGKDTNTGNSFSDSPVSHPESDAAI 810
 Db 2155 -----GEGLGRNLPPADRNLTNLSDEEBDDDDDEVOEETPPSEAEGEHV 2203
 Qy 811 NVEK-----LSGDES----- 820
 Db 2204 ETEBETKPVKTEGAGATEVTKQGSAPTATPTTVEDICATVAKALGDKSLNAACALKY 2263

QY 821 ----- 820
 Db 2264 GKNSRLGWKIPSGDKTTSSENGAPRRASAHGKSDSEKSGICVPPRRRLYIKKIV 2323
 QY 821 ---SSETRGILDIN-DPSVTNNVNEHVDASNTQGSVNT-----SDITNGHSESS----- 866
 Db 2324 DWAESQKTVTSVNGDNGQGVSVNGASESGSGSGTSESQADSVSQNGASTSPOVAL 2383
 QY 867 LNRITNAQDI-----KIGRSGNQSDNQENSSH-----SSDNGSLTIGQVPSSED 911
 Db 2384 LHAFVKSAAIETFPFAWHKYVDKE-IEEKEQAAQHLVQRKTSNPQKLEGEIP-ED 2441
 QY 912 NTQNTYDSQPHRD-----TPNALASLPDDKINET-EGFDSRDSEN-----GRGDTTS 960
 Db 2442 FKQMFYTLGDYRDLVGDKTMIETAELEKSGDGTIEDISEKIPKILDGNNKAAGGPKQP 2501
 QY 961 NTHDVRRTNIVSERRVNSHDFIRNGM-----ANNNAHQYITOI-----ENNGIIRQEE 1010
 Db 2502 NSGKTPQ-----EWWKENAKHIWHGICALTYNTDSNGKDKKIQQVKATNTDLFQKLKK 2556
 QY 1011 -----SAGNSVNYKDNPKRNSFSSNDHKNIQYNSRDTKRVRBEEIIKLSKQNK 1060
 Db 2557 DNDYETVSFGASGTGAKSNDOTKLKNPVVPTFYFRWLEEWGEFCRQKHLYIKKDCR 2616
 QY 1061 CNNEYSMEYCTYSDERNSSPGCSREERKKL-CCQISDYCLKYFNFYISYIYNCIKSEIK 1119
 Db 2617 DN-----KFCSGDGLRCDKVPDKKIDFKHFDPCSCARHCRSYRKWT-----ERK 2661
 QY 1120 SPEYKCFKSGQSSIPYFAAGILVIVILLSSASRMGKN-----BEYDIGESN----- 1169
 Db 2662 KTEVE--KQE-----SAYSKQKSNVYNGSGDGNNDKEFYT 2697
 QY 1170 -IEATFEENNYLNKL 1183
 Db 2698 KLETCTRATNFLESJ 2712
 RESULT 31
 AEB22174
 ID AEB22174 standard; protein; 3553 AA.
 AC AEB22174;
 XX
 XX 22-SEP-2005 (first entry)
 XX
 XX Plasmodium falciparum VAR4.
 XX
 XX antimalarial; vaccine; protein purification; pharmaceutical; diagnostic;
 KW Plasmodium falciparum infection; infection; VAR4.
 XX
 XX Plasmodium falciparum.
 XX
 XX WO2005063804-A1.
 XX
 XX 14-JUL-2005.
 XX
 XX 28-DEC-2004; 2004WO-DK000922.
 XX
 XX 30-DEC-2003; 2003DK-00001954.
 XX
 XX (UYKO-) UNIV KOBENHAVNS PANUM.
 XX
 XX Jensen ATR, Hviid L, Jorgensen L, Lavsten T, Magistrado P;
 PI Nielsen MA, Salanti A, Staaleo T, Theander TG;
 XX
 XX WPI; 2005-522559/53.
 DR N-PSDB; AEB22173.
 XX
 XX New isolated VAR4, VAR5, and VAR6 polypeptides isolated from Plasmodium
 PT falciparum, useful as vaccines for treating or preventing severe malaria.
 XX
 XX Claim 1; SEQ ID NO 2; 193pp; English.
 PS
 XX

CC The invention describes an isolated polypeptide comprising an amino acid
 CC sequence selected from at least one of SEQ ID Nos 2, 4, and 6 (3553, 2994
 CC and 3346 amino acids) for use as a medicament. The polypeptides, nucleic
 CC acids and vectors are useful as vaccines for preventing and/or treating
 CC malaria. The malaria is preferably severe malaria caused by Plasmodium
 CC falciparum. An in vitro diagnostic method is useful for detecting anti-
 CC malarial antibodies. The polypeptides are useful for generating a vaccine
 CC against severe malaria, where the vaccine comprises antibodies generated
 CC against the polypeptide which have cross-reactivity to parasites causing
 CC severe malaria and the ability to inhibit adhesion to endothelial cells.
 CC The polypeptides are also useful in screening for inhibitor molecules
 CC capable of inhibiting binding of any of the polypeptides to a receptor
 CC expressed on endothelial cells. This is the amino acid sequence of
 CC Plasmodium falciparum Var4 that is transcriptionally upregulated in
 CC Plasmodium following increased antibody recognition.
 XX
 SQ Sequence 3553 AA;
 Query Match 9.1%; Score 592; DB 9; Length 3553;
 Best Local Similarity 20.0%; Pred. No. 2.8e-27;
 Matches 315; Conservative 205; Mismatches 511; Indels 544; Gaps 68;
 QY 39 SDISRVNSBELNNHKTNIYDSYEDVNNKLINSFVENKSVKKRSLSPINNKTYSYDII 98
 Db 1252 TDVAKILOGEAN-----ETMLKNSNGNDKESKLKGAESGD-- 1289
 QY 99 PPSYSYRNDKFNLSLSENDNSGNTNSNFANTSEISIGKDNKQYTFIQKRTLHFAAGIKR 158
 Db 1290 ---YS-RGGTSPDFNNLGGITQKSHANDSQPCYCKDQKFNV----- 1331
 QY 159 KSIKWICRNSSEKIT---VCVPDRKTQLCVANFLNSRLTM--EKPEKIP- 203
 Db 1332 -GTSEFKDNHRKRTPEAYMPPRREHICTSN-----LEYLIHKRKKPIEGDPNKIHS 1385
 QY 204 -----LISVNTAE--KLINYKNEG-KOPSIFCNELRNSPSPRSFIDGDMDFGGWTD 254
 Db 1386 LIGDVLAAKYBAENIKLYBENNRRKQDEGICRAMKYSFADIGDIIRGKDM-WIENDDA 1444
 QY 255 VKGYIN-----TKFSD-----YKKEKNVEKLANIKKEWKEKNKANLWNHNVNKG-NIS 303
 Db 1445 KRLQTNLKEIFTYKIEKTGTYINEDNDPYL-KLRADWEANRAKWKAMKCKTNGVDIT 1503
 QY 304 KECALIPABEE--PQINLWIKENENFLMEKKRLFLMIKQCVB-----NKKYEACF----- 352
 Db 1504 CDSHTPLDDYIPQRLRWMTAEAWYCKAQSOEYKLEBKCSQCKSGKGGNCEYRETKE 1563
 QY 353 -GGCLPCCSSYTSFPMKSKTOMEVLTN---LYKKNSGV-----DGN--NFLN 394
 Db 1564 CNDCKQACEYRKIKTWDQWKVISNKYEDLYKKAQNPNTNAVLRKDKOEKKNVIDFLT 1623
 QY 395 DLFPKKN-----KNLDD-----FPKNEKE-Y 415
 Db 1624 QLOKANNKEKTVHTVYSTAAGYIHQEARTECOEORFCDKNGKIDNTSYAFKDPHGY 1683
 QY 416 DDLCDRCRYTATIKSPFNGPAKNDVDIASQINV-----NDLRFG---CNYK-----S 460
 Db 1684 ATACDC-----INRSOTEPEPKKEENVEACKIVEEVLSPKPRKTTGGIDHCNPKYPRK 1738
 QY 461 NNEKSNCT-GTFTNKFPCTCEPPRQTLCLGRTYLLHRGHEEDYEKHLGLASIVEAQL 518
 Db 1739 ENYFGWNCPTGQFKSGHAGACMPPRRIKLCVINQLYNLEKKSPEELRKAFIGCAAIETYW 1798
 QY 519 LKYKYEKEDENALCSTIONS-----YADLADIIGSDIIXDYDGK---KM 560
 Db 1799 LMQYKKKXNGGVAQAKLNSGTIPDDFKQMFYFGDYRDLCLDTJSSKADTSTGVGVK 1858
 QY 561 EENLNKV-NKDKRNEESLKIIFREKWDENKENVKNSAVLKNKETCK--DYDKFOKI- 616
 Db 1859 KINIDSVFQKIDITNVEQ---RKPWWGKNAEAIWDMCLCALSNTTNKMDYNAHTKLN 1914
 QY 617 -----PQFLRWFKWGDGDFCEKKEKTIYSPESFKVECK----- 649
 Db 1915 PTYGVNAIKSELEDVFNRFQFLWFTWNSDEFCTERSIKIKELETCKNDCVTSESSTDA 1974

QY 650 --KKDCDE---NTCKNCSBYKKWIDLKSEYKQVDKYTKDKNKKWYDNDIV-KNKE 702
 Db 1975 TGNKTCDDKCKDECKRATYTKWLKNWKYQYTKQSKYFDDKRLKYKSIDVASSTQ 2034
 QY 703 ANVYLKESK-----ECKVNFDDKIPNESP-----NEYEDCKKC----- 738
 Db 2035 AYQYLHAQLKLCGNADCKMDGSKETTGQPDNSHSHMPASLDDEPEEVNGKCNCKVK 2094
 QY 739 -----DEIKYLN-----EIKYKTKHD 755
 Db 2095 HROPPLALPPAPSGPPAEQIEHNRGRSERGQGPLPARPPPPQAAQPPQPKPRT 2154
 QY 756 IYDIDTSDFGDK-----TPISINANINEQQSGKDTNTGNSSETSPVSHESPESAAI 810
 Db 2155 -----GEGLRNLPADRNTNLSDEEDEDDEVEQEEETPPSEAGEGSHV 2203
 QY 811 NVEK-----LSGES----- 820
 Db 2204 ETEBETPKVKEKTEGAGATEVTKQGSAPTATTPTVEDICATVAKALGKDSLNAACALKY 2263
 QY 821 ----- 820
 Db 2264 GKNSRLGKWKIPTSGDKTDTSENGAPRRASAHGKSDSEKSGICVPPRRRLYIKIV 2323
 QY 821 ---SSETRGILDIN-DPSVTNNVNEVHDASNTQGSVNT---SDITNGHSESS----- 866
 Db 2324 DWASQSKVTIVSVDGNGSEVSVNGASGSGSGTESQADSVSQNGASTSPQVAL 2383
 QY 867 LNRTTNAODI-----KIGRSGNEQSDNQENSSH-----SDNSGSLTIQGVPSBD 911
 Db 2384 LHAFAKSAAIETFFAWHKYKVDKE-IEEKEQAAQNLVORKTSENPOKLEGGEIP-ED 2441
 QY 912 NTQNTYDSQPHRD-----TPNALASLPDDKINEI-EGFDSRDSSEN-----GRGDTTS 960
 Db 2442 FKQMFYTLGDRDILVGDKTMI EALEKSGDGTIEDISEKIPKILDGNNKAAGGPKQP 2501
 QY 961 NTHDVRRTNIVSRRVNSHDFIRGM-----ANNAHGOYITQI---ENNGIIRGOEE 1010
 Db 2502 NSGKTPQ-----EWWKENAKHIWGMICALTYNTDSNGKDKKIQOVKATDNTDLFOKLKX 2556
 QY 1011 -----SAGNSVNYKNDPKRNSFSENHDKNIOEYNSRPTKRVREIILKSKQNK 1060
 Db 2557 DNDYETVSFGASGTGAKSNDTDLKNFVVRPTFYRWLEEWGEFCRQKQKHLIYIKKDCR 2616
 QY 1061 CNNEYSMEYCTYSBERNSSPGCSREERKLI-CQOISDYCLKYFNFYFIEYNYCKSEIK 1119
 Db 2617 DN-----KFCSGDGLRCDEKVPDKDIFKHPDCPSCARHCRSYRKWI-----ERK 2661
 QY 1120 SPEYKCFKSGQSSIPYFAGGILVIVLLLSASRMGKSN-----EEYDIGESN----- 1169
 Db 2662 KTEVB--KQE-----SAYSKQSNVNGSGDGGNNNDKEFPYT 2697
 QY 1170 -IEATFEENNYLNLK 1183
 Db 2698 KLETCTKATNFLES 2712

RESULT 32

AEB22176
 ID AEB22176 standard; protein; 2994 AA.
 XX AC AEB22176;
 XX DT 22-SEP-2005 (first entry)
 XX DE Plasmodium falciparum VAR5.
 XX KW antimalarial; vaccine; protein purification; pharmaceutical; diagnostic;
 XX KW plasmodium falciparum infection; infection; VAR5.
 XX OS Plasmodium falciparum.
 XX

FN WO2005063804-A1.
 XX 14-JUL-2005.
 XX 28-DEC-2004; 2004WO-DK000922.
 XX 30-DEC-2003; 2003DK-00001954.
 XX (UYKO-) UNIV KOBENHAVNS PANUM.
 XX Jensen ATR, Hviid L, Jorgensen L, Lavsten T, Magistrado P;
 XX Nielsen MA, Salanti A, Staalso T, Theander TG;
 XX WPI; 2005-522559/53.
 XX DR N-PSDB; AEB22175.
 XX PT New isolated VAR4, VAR5, and VAR6 polypeptides isolated from Plasmodium
 XX falciparum, useful as vaccines for treating or preventing severe malaria.
 XX PS Claim 1; SEQ ID NO 4; 193pp; English.
 XX CC The invention describes an isolated polypeptide comprising an amino acid
 CC sequence selected from at least one of SEQ ID NOS 2, 4, and 6 (3553, 2994
 CC and 3346 amino acids) for use as a medicament. The polypeptides, nucleic
 CC acids and vectors are useful as vaccines for preventing and/or treating
 CC malaria. The malaria is preferably severe malaria caused by Plasmodium
 CC falciparum. An in vitro diagnostic method is useful for detecting anti-
 CC malarial antibodies. The polypeptides are useful for generating a vaccine
 CC against severe malaria, where the vaccine comprises antibodies generated
 CC against the polypeptide which have cross-reactivity to parasites causing
 CC severe malaria and the ability to inhibit adhesion to endothelial cells.
 CC The polypeptides are also useful in screening for inhibitor molecules
 CC capable of inhibiting binding of any of the polypeptides to a receptor
 CC expressed on endothelial cells. This is the amino acid sequence of
 CC Plasmodium falciparum Var5 that is transcriptionally upregulated in
 CC Plasmodium following increased antibody recognition.
 XX SQ Sequence 2994 AA;
 Query Match 8.6%; Score 554.5; DB 9; Length 2994;
 Best Local Similarity 21.4%; Pred. No. 4.9e-25;
 Matches 286; Conservative 172; Mismatches 425; Indels 451; Gaps 65;
 QY 114 ENEDNSGNTNSNPNANTSEISIGKDNKQYTFIQKTHLFA---CGIKRSIKWICRENSE 170
 Db 839 KEEDTDGKQPPPAATTPGVKPCDIVEKFKDKHDTGADHCNPKKDYPPW---KNDK 895
 QY 171 KIT-----VCVPDRKIQLCVANFLNSRLTMEKFEFLISVNTBAKLLYNK----- 217
 Db 896 SLVDEGCVMPPRQKLCVINLHFHFKENTSDLLREAFIKCAAETYLWQYKEDNNGE 955
 QY 218 -----NEGKPSIFCNELNSPSDFRSSFIDGMD-FGGNTDRVKGYINTKPSDYKEK 270
 Db 956 DLQNLKSGKIPEDFKRQMFYTFGDRDPLFGTDISKLNKHTKTEAVKTNIDRIFPP----- 1010
 QY 271 NVEKLNN-IKKEWKKNAKLNWNHMI--VNHKGNISKECALIP----- 310
 Db 1011 -TERTNDTIRKEFWKNAESIWOGLCALSTNSNDKKDDVQKELNSTYNTDIKNLE 1069
 QY 311 --AEEPOINLWIKENENFLMEKRLFLNIDKCVENKTYEACF-----GGCRLPCSS 361
 Db 1070 DFANRPQFLRFIEMSDFCRERKKEKVGSA--KNDYEGCANTKDNNGNCVACNA 1127
 QY 362 YTSFMKSKTKQMEVLTNLY---KKKN-----SGVDKNNFLND-----LFFKN 401
 Db 1128 YKYYITDKGEYKQAKKFDIDKSNKPGYEDYSGKASELYLKEKINCINSSCDYMLKLDN 1187
 QY 402 KNDLDDFFKKEKEDD-----LCDCRYTATII--KSFINGPAKNDVIDIASQINVDLRGP 454
 Db 1188 SNWVE---KPHTTYDDNSLQNKSCPLSPCEIVDTLGDTKTSKYAE----- 1231
 QY 455 GCNYKSNNEK---SWNCTGT--FTNKPFGTCPPRRQTLCLIG--RTYLH----- 497

Db 1232 GCKWKYKMPGLGLWLCNDKEGKEDGLCLIPRRKRLVYKDLTSPDHTTGLREAFI 1291
Qy 498 -----RGHEDYKHL-----LGASIIYEAQLKYKKEKDENALCSIIQN 537
Db 1292 KCAAVETFAWHEFTKEKREYKEKQNGELGFIENDQIPK-----DPDNPNQKIRKN 1346
Qy 538 -----SYADLADILIGSDI-IDYVGKMEENLNKVNKDKK-RNESLKIPIR 582
Db 1347 GEIHEBFKSQMFYTLADRIILFNNIGIGNDMG-KVKSNDIKVFANSSGKTPTAKTTP 1405
Qy 583 EKWDENKENVWVMSAVL-----KNKETCK-----DND 611
Db 1406 KEWKEKNAKDIWEGMLCALSYDTTKIKNEELRKLIDPKSNVWYKVTFFSSDNTNLS 1465
Qy 612 KFOKIPOLRWPKEGDDCEKREKIKYSPESFVCK-----KKOCD-----E 655
Db 1466 KFTERRPPFRWFQWGEFCKRKKIKIDKIEK-----ECRGPVGRNHCDGDGDCSEIGPNE 1522
Qy 656 N-----TKNKKSEYKWKIDLKSEYEKQVDKYTKD--KNKKMYDNI---DEVKN- 700
Db 1523 NGSPAIFKPCSAISCRSYKTWINTKKDEFKQKLYNKIKONKSNYDNIYDKFVKNL 1582
Qy 701 ----KEANVYLKEK-----SKECKOVNFDDKIFNES--PNEYEDMC-----735
Db 1583 CTDYKSVDSFLKLKEGPCCKNKTDSKIDFKOTEETFRNAEYCDPCPVGVICNNGDCS 1642
Qy 736 ----KKCD--EIKYLNBI-----KYP-----750
Db 1643 NSTEKKCDAEFKYVDVKNKENPKNVNLVSDKTAKYPGLDNGVCENSSIFEGIRE 1702
Qy 751 -----TKKHDIV-----IDTF---SDTFGGTPIISINANI 778
Db 1703 KWSGCGYCGLDICPNKTTGDIHDKQNAPIRVLPKRWIENFLKDHKKIKDLSLCIN--- 1759
Qy 779 NEQSGKSDTNTNGSE-----TSDSPVSH-----EPE 805
Db 1760 NENRNICDVCRNKCECIDKWIEMKEMKIVRDYVYKQYNVADSVYVYRFLGLOPQ 1819
Qy 806 SDAAINVEKLSGDESSETRIGILDINDPSVNNVNEVHDASNTQGSVSNSTSDITNGHSE 865
Db 1820 ND-----LEKVG-----DVNDLRDLLELSECTNTVSTENRCKRKKVV---ES 1860
Qy 866 SLARTTNAQDIKIGRSGNEQSDNCSHSDNSGSLTIGQVPSD---NTQNTYDSQN-P 922
Db 1861 LLNLKLN-----EIRHCKNERDDSMGKES-----CKTLPE-PTDDPQTDSDTHDPDIP 1908
Qy 923 HRDTPNALASPSDDKINEIEGFDSSRDSENGRGTTSNTHDVRRTIVISERRVNSH-DF 981
Db 1909 PGDVAPTFCNVPA-----NPGCD-KSATNVNVNTEVAKEMHEEAHKOM 1950
Qy 982 IRNGM--ANNNAHHQYITQIENNGIIRGOBESAGSNVYKDNPKRSNFSSENHKKVIOE 1039
Db 1951 LERSVKVSKVDSTVESVLRADASKGYKHGEGNPDLDKHNMCNITKEHTNYQKRGYN 2010
Qy 1040 Y-----NSRDTKRV-----RBEILKL-----SKQCNKNEYSMEYCTYSDERN 1077
Db 2011 YRGCTKGNGKOTRFVIGTIWKEDEKDEKTIKVLPLPRRHMCST--NLEYLLHVNK-- 2066
Qy 1078 SSPGPCSREERKKL 1091
Db 2067 ---GFLKVPDKI 2077

RESULT 33

ABB07655
ID ABB07655 standard; protein; 302 AA.

XX ABB07655;

AC ABB07655;

XX 20-MAY-2002 (first entry)

XX P. falciparum EBA-175 protein PfP1 domain sequence.

XX Duffy binding-like domain; DBL domain; parasite; malaria; DBP protein;
XX protozoacide; vaccine; EBA-175 protein.
XX Plasmodium falciparum.
XX WO200212292-A2.
XX 14-FEB-2002.
XX 03-AUG-2001; 2001WO-BP009023.
XX 07-AUG-2000; 2000GB-00019375.
XX (ITGE-) INT CENT GENETIC ENG & BIOTECHNOLOGY.
XX Chitnis C, Pandey K, Singh S, Yazdani SS, Pattnaik P;
XX WPI; 2002-227139/28.
XX Producing polypeptide with Duffy binding-like domain, by expressing
XX polypeptide in bacterium/yeast, extracting and denaturing it, refolding
XX polypeptide in presence of arginine and urea, and optionally recovery.
XX Disclosure; Page 43-44; 47pp; English.
XX The invention relates to a method for producing a polypeptide (I)
XX comprising a Duffy binding-like (DBL) domain. The method involves
XX expressing (I) in a bacterium, or as a non-secreted polypeptide in a
XX yeast, extracting the expressed polypeptide from the bacterium or yeast
XX and denaturing the polypeptide, refolding the extracted polypeptide in
XX the presence of arginine and urea, and optionally recovering the refolded
XX polypeptide. The method is useful for producing a polypeptide comprising
XX DBL. (I) is useful for identifying a substance that modulates the
XX interaction between the polypeptide and a host cell receptor involved in
XX the entry of a parasite into a host cell. The substance identified is
XX useful in the manufacture of a medicament for treating or preventing
XX malaria. A pharmaceutical composition or a vaccine composition obtained
XX by formulating the refolded polypeptide is useful for treating or
XX preventing malaria. (I) is useful as vaccine to prevent malaria or
XX infection by P. falciparum or P. vivax. The present sequence represents
XX the P. falciparum EBA-175 protein PfP1 domain sequence
XX Sequence 302 AA;
Query Match 8.4%; Score 546.5; DB 5; Length 302;
Best Local Similarity 37.7%; Pred. No. 6.5e-26;
Matches 122; Conservative 50; Mismatches 119; Indels 33; Gaps 9;
Qy 121 NTNSNFPANTSEISIGKONKQYTFIOKRTLPACGIRKRSIKWICRENSKIT-VCVPDR 179
Db 3 NTSSNN-----EV-----LGNCREKRGKMKWDCKKNDRSNYVCIPDR 40
Qy 180 KIQLCVANFLNRLTMEKFKBIFLISVNTAEKLLYNKNEGKDPISFCNELRNSFSDFS 239
Db 41 RIQLCVINLSIITYTKETMKDHFTEASKESQLLLKNDKNKYNKFCNDLKNKSLDYGH 100
Qy 240 SFTGDDMDPGGNTDRVKGYNITKFSDYKKNVKEKLNNIKKEWKEKNKANLWNHMIWNH 299
Db 101 LAMGNDMDFGYSTKAENKIQEVFKGAHGEISEHKIKNFRKEWNEFREKLWEAMLSEHK 160
Qy 300 GNISKECALIPAEPOINLWIKENENFLMEKRLFLNIDKCVENKCYACGGRLPC 359
Db 161 NNIN-NCKNIPQBELQITQWIKWEHGEFLERDNRSLPKSKCNNTLYEACEKCEIDPC 219
Qy 360 SSYTSFMKSKTQMEVLTNLYKKNSGVDKNNFLNDFK-KNNKND--LDDFFKN-EKEY 415
Db 220 MKYRDWIIIRSKFEWHTLSKEYETQK--VPKNAENYLIKISEKNDAKVSLLNNDAEY 277
Qy 416 DLDCCRYTATIKSFINGPAKND 439
Db 278 SKYCDCKHTTILVKSVLNG---ND 298

RESULT 34

AAW00384
ID AAW00384 standard; protein; 2913 AA.

AC AAW00384;

DT 16-OCT-2003 (revised)
DT 21-FEB-1997 (first entry)

XX Plasmodium falciparum erythrocyte membrane protein.

XX Plasmodium falciparum; erythrocyte membrane protein; malaria; detection;
KW identification; treatment; prevention; parasite.

XX Plasmodium falciparum; MC type.

XX Key Location/Qualifiers

FT Domain 62..394

FT Region /label= Duffy binding ligand domain 1

FT Region /note= "Cysteine rich motif"

FT Domain 839..1272

FT Region /label= Duffy binding ligand domain 2

FT Region /note= "Cysteine rich motif"

FT Domain 1706..2005

FT Region /label= Duffy binding ligand 3

FT Region /label= Duffy binding ligand 4

FT Region /note= "Cysteine rich motif"

FT Domain 2450..2475

FT /note= "Putative transmembrane domain"

XX W09633736-A1.

XX 31-OCT-1996.

XX 26-APR-1996; 96WO-US005798.

XX 27-APR-1995; 95US-00430908.

XX (APFY-) APFYMAX TECHNOLOGIES NV.

XX Baruch DI, Pasloske BL, Howard RJ;

XX WPI; 1996-497376/49.

XX N-PSDB; AAT41852.

XX New Plasmodium falciparum erythrocyte membrane proteins - used to develop

XX products for the diagnosis, treatment or prevention of malaria parasite

XX infections.

XX Claim 1; Fig 12; 149pp; English.

XX A polypeptide comprising a Plasmodium falciparum (Pf) erythrocyte

XX membrane protein 1 (PfEMP1) or active fragments or analogues of that

XX protein can be used in the treatment or prevention of symptoms of a

XX malaria parasite infection. The polypeptides can inhibit, block or

XX reverse the sequestration of erythrocytes in patients suffering from

XX malaria. Nucleic acids derived from the PfEMP1 gene can be used as probes

XX and primers to identify a Plasmodium falciparum parasite, the primers

XX used to generate characteristic amplification patterns from different P.

XX falciparum strains. Antibodies specifically immunoreactive with the

XX PfEMP1 polypeptide or its fragments may be used in diagnosis of malaria

XX infection. This is the PfEMP1 protein of the MC type of Plasmodium

XX falciparum. An alternative, truncated PfEMP1 protein is given in

XX AAW00385. (Updated on 16-OCT-2003 to standardise OS field)

XX Sequence 2913 AA;

XX Query Match 8.4%; Score 546.5; DB 2; Length 2913;

XX Best Local Similarity 22.6%; Pred. No. 1.5e-24;

Matches	279; Conservative	182; Mismatches	435; Indels	337; Gaps	65;
Qy	48 ELANNHKTNIYDSYEDVNNKGLINSFVENKSVKKRSLSFINNKTYSYDIIPPSVSYEND	107			
Db	1663 KINNELKGN-----GKDFNGKCNN-----VKKNGAVI-----GEESC	1695			
Qy	108 KFNLSLEEDNSGNTNSNFANTSEISIGKDNKOYTF-IQKRTHLFAGGIRKRSIKWICR	166			
Db	1696 KFEQTYENSYN-----NINNK-----KDNQNERFKIGQKWNFKYGTIRKDL-----	1738			
Qy	167 ENSEKITVCPDRKIQLCVANF-----LNSRLTWEKFEIIFLISVNTAKLLYNQVE	219			
Db	1739 -----CIRPREHMLDLSMLGRTTISDSSALLKIOEAAKSERDDIIRKLEQNS	1790			
Qy	220 GKPSIFCNELRNSFDRSSFIGD--DMDFGGNTDRVKGY--INTKFSDYK-----	268			
Db	1791 CDEHRI-CDAMKYSFAD-----LGDIIIRGRDLWNKSKQGLQKRLEYAFINYNKLQND	1844			
Qy	269 -----EKNVEKLNINIKCEWWEKANKANLWNHMTVN-----HKG-----	300			
Db	1845 KNYEKORPKYQLQRLSDWDANRKHINWAMTCNAPDADAKFLKKPNPDTSGSSSGKIMWT	1904			
Qy	301 ----NISKCAIIPAE-----PQINLWIKENWENF-----LMEKKRLFLNLIKQCVN-----	345			
Db	1905 HSNCGYDKB-----PPDYDIYIPQPPRMQWSESPCKLLNEEMEQFEKTCGCKKNSITCE	1960			
Qy	346 --KKYEAFCGGCLPCSSYTSFMKCKSK-----TQMEVLTNLYKKNSGVDKQNFNDLPFK	399			
Db	1961 DRRNGTNC-ENCRNQCEKYKGLIHNWKLGFQYKYEIYNNKDSKINSNEYFKFLJEK	2019			
Qy	400 -----NNKNDLDDPFKN-EKEYDDLCDC-----	421			
Db	2020 LKDKCKELNSDKCIDETHCTKYKFSNSENKNNNYAFKPPKEYEKACKCDADPDLN	2079			
Qy	422 --RYTATIISFLNGPAKNDVDIASQINVDLRFQGCNYK--SNNEKSWNCTGFTTNKFP	477			
Db	2080 CPKDSATVEK-----ACNTL-LPTKL-----CESKTFNDDSDWT--SFVQTSF	2121			
Qy	478 ----GTCEPPRRQTLCLGRTYLLHRGHE--EDYKEHLGLGASIIYEAQLLYKYKEDNAL	531			
Db	2122 RDNTGLVPPRRRQICLNITTKLRSIEKIDDFKAEMLTSAYNEGKLLCELYK-KDRDVT	2180			
Qy	532 CSIIQNSYADLADIIKGSDDIKDYGGKMBENLNKVKDKKRNEESLKIPEKRWENKE	591			
Db	2181 LQWKYSFYDGDIVKGTDLIISTAPLDKLTKNLVLLKGDGTNE--IKEDRGKWWTEWRT	2238			
Qy	592 NVWKVMSAVLK-----NKETCKDYKFKQIPQFLRFWKWGGDDFCERKEKIKYSFESFK	645			
Db	2239 RVWHAMLCGYKAAGGKIEERDCSLPD--DNTHQFLRWFWSEHFCAKQK--LFNEVK	2293			
Qy	646 VECKKDC-----DENTCKNCKSEYKWKIDLKSEY-----BKQVD-----	681			
Db	2294 RECASAQCIIEYGTIDPPVCEEAQTYRDYITRKIQBYRLNYQYNTNFNEKKAETVKAP	2353			
Qy	682 KYTKDKN-----KKWVDNID--EVKNKEANVYLKESKECKDQVDFDK	722			
Db	2354 EYFNKCNKCNKCLSKYIDIEKKWKNNYDSFDDNLDKKNKICRQIKPK-RPPKVKPKEE	2412			
Qy	723 IFNESPNEYE---DMCKCKDBIYKLYNEIKYKPKTKHDIYDIDTFSDTFGDGPPISINAMIN	779			
Db	2413 ---HTPSEQDTPLPLPKPDDL-----PPPAEPPNRDILEKTIYPPGIALALGSIAF	2461			
Qy	780 EQSGKDTSTNGSETSDSPVSHPEPESDAAINVEKLGSDE-----SSSETRGILINDPSV	835			
Db	2462 LFLKCKTKSSVGNL---FQILHIPKSDYDIPT-KLSPNRYPYPTSGKYRG-----	2507			
Qy	836 TNNVNEVHDASNTQGSVNTSDITNG-----HSESSLN-----RTNAQDIKIRSGN	883			
Db	2508 KRYILEGDSGTSGSYTDHYSDITSSSESEYEEEDINDIYVPGSPKYKTLIEVLEPSEN	2567			
Qy	884 EQSDNQNSHSHSDNSGLTIGQVPSBDNTQNTYDS-----ONPHRDTNPALA--	931			
Db	2568 NTHASGKNT--PSDTQNDIQNDGIPSSKITDNEWNTLKDEPISNMLQNEPTEPNMLGYN	2625			

QY 932 ----SLPSDDKIN-EIEGPDSS---RDSNGRGDTTNTDHDVTRTIVISERRVN--SHDF 981
Db 2626 VDNHTPTTTRHRHVEEKPFMSIHDRLYSGE-EYSYNVMVNDTPIISARNGNYGGIDL 2684
QY 982 IRNGMANN--NAHQYITQIENN--GIIRQERSAGNSV--NYKDNPKRSN---FSGEND 1032
Db 2685 INDSLNSKVDIYDELLAKKENELFGTNHTTKNTSTNSVAKTNTDPIHNQLNLFHTWLD 2744
QY 1033 HKKNIOEYNSRDYKRVREEII-KLSQK-NKCNN 1063
Db 2745 RHRDMCE--KWDNTNNKKEILLKKBWNNKDN 2775

RESULT 35

AE822178
ID AEB22178 standard; protein; 3346 AA.

XX AC AEB22178;

DT 22-SEP-2005 (first entry)

XX Plasmodium falciparum VAR6.

XX antimalarial; vaccine; protein purification; pharmaceutical; diagnostic;
KW plasmodium falciparum infection; infection; VAR6.

XX Plasmodium falciparum.

OS WO2005063804-A1.

PN 14-JUL-2005.

XX 28-DEC-2004; 2004WO-DK000922.

PF 30-DEC-2003; 2003DK-00001954.

XX (UYKO-) UNIV KOBENHAVNS PANUM.

PI Jensen ATR, Hviid L, Jorgensen L, Lavsten T, Magistrado P;

PI Nielsen MA, Salanti A, Staalso T, Theander TG;

XX WPI; 2005-522559/53.

DR N-PSDB; AEB22177.

XX New isolated VAR4, VAR5, and VAR6 polypeptides isolated from Plasmodium
PT falciparum, useful as vaccines for treating or preventing severe malaria.

XX Claim 1; SEQ ID NO 6; 193pp; English.

XX The invention describes an isolated polypeptide comprising an amino acid
CC sequence selected from at least one of SEQ ID Nos 2, 4, and 6 (3553, 2994
CC and 3346 amino acids) for use as a medicament. The polypeptides, nucleic
CC acids and vectors are useful as vaccines for preventing and/or treating
CC malaria. The malaria is preferably severe malaria caused by Plasmodium
CC falciparum. An in vitro diagnostic method is useful for detecting anti-
CC malarial antibodies. The polypeptides are useful for generating a vaccine
CC against severe malaria, where the vaccine comprises antibodies generated
CC against the polypeptide which have cross-reactivity to parasites causing
CC severe malaria and the ability to inhibit adhesion to endothelial cells.
CC The polypeptides are also useful in screening for inhibitor molecules
CC capable of inhibiting binding of any of the polypeptides to a receptor
CC expressed on endothelial cells. This is the amino acid sequence of
CC Plasmodium falciparum Var6 that is transcriptionally upregulated in
CC Plasmodium following increased antibody recognition.

XX Sequence 3346 AA;

Query Match 8.4%; Score 543.5; DB 9; Length 3346;

Best Local Similarity 20.8%; Pred. No. 2.8e-24;

Matches 300; Conservative 193; Mismatches 545; Indels 403; Gaps 64;

QY 31 LYNQDESSD---ISRWNSP-----ELNNHKTNIYDSYEDVNNKL 69

Db 730 LLNDEDEDDDETTPRAHNPCTVDKNDSPQTKTVSYIARQMRRAKQMTKNSVVDGNKL 789
QY 70 INSFVENKSVKKRSLSFIN--NKTYSYDIIIPPSYSYRNDKFNLSLENEDNSNTNSNN 126
Db 790 EGDIPK-----VTFRGGVGKNLGDIKIDKTSND-----SRGPTDGPCGKG 835
QY 127 FANTSISIGKD-----NKQYTFI---QKRTHLPACGIKRKSIIKWI CHRENSSEKIVCV 176
Db 836 ----DRFKIGTDMQGDSPVNPQYRGIYMPRROH-----FCTSNLEKLDVSR 878
QY 177 PDKIQLCVANFLNSRLETMEKFEFLISVNTAEKLLYNK--NEGDPGSI FNCNELNSF 234
Db 879 VTRN-----GNASNSLLGDV-----LLAAYEABRTKNHYVSKKEHSEACRAVYSF 926
QY 235 SDFRSSFIGDDM---DFG-----GNTDRV-----KGVINTKFSDDYK--EKNVEKLNLIK 279
Db 927 ADLGDIIIRGDMWDKNHGEKKTQENLERIFAKIKEQLLNSSIKDKYKDDKATPKYKQLR 986
QY 280 KEWWEKXKANLWNHMIWNHK-GNISKECALIIPABE--PQINLMIKEWNEFLMEKKRLFL 336
Db 987 EDWWEANRSQVWEAMQCPKNGTFFCKSDHTPLHDYIPQRLRWMTAEAWYCKEQSRLYG 1046
QY 337 NTKDKCVENKKYEACFGG-----CRLPCSSYSFNMKSKTKTOMEVLTNLY-----KKK 383
Db 1047 ELVETCGKMHGKCKQGNHCHVTCRPAKEYKFFINTWQPQWQMEQKYSQLYEAKKY 1106
QY 384 NSGVDXN-----NFLNDLFKNNKNDLDD-----407
Db 1107 NDSRRKDTTKDDYVLQFLNKLTLQNGKNKYDTYTAEGYVHQEAHISDCQKQTOPCKKRG 1166
QY 408 -----PPRNEKEYDDLDCD--RYTATIIKGFNLGPAKNVDVDAQ-----IN 447
Db 1167 EIPSSDTETDNNYAFRPPQPHDHDVECECNTRQTKVRKK-----KKKVDACEMAKTLH 1220
QY 448 VND--LRGFGCNYSNNE--KSWNCTGTFTNKPGCTCPRPRTQTLCLGRTYLLHRGHE- 501
Db 1221 NNDGTIRIGQCKRCKDEGNAEYPKWDCNSQIHHTHNGACMPRRKQLCV---YFANPSQI 1277
QY 502 --EDYKEHLLGASIEAQLLKVK-----YKEKD--ENALCSIIQN-----SYAD 541
Db 1278 GSINKQDNLKAFIISAAAEFRSQYKSKNGENLQTLKDGTPIDDFKROMFYTYGD 1337
QY 542 LADIHKGSIIIDY-YGKQWENLKNVKNKQKR---NEESLKI FRKXWDENKENVKVM 597
Db 1338 YRDLFGTIDSKLGEGTALRQINILFPNGVRKIPNEKT---REKWWTDHGPEIWKGM 1393
QY 598 SAVLKN-----KET--CKDY-----DKFOKIPQFLRWFKWGDGDPCEKR 634
Db 1394 LCAITNGLSESEKTKIFDDYSHDKVQSKNGNPSUEDFAKQPFRWFLEWSEDFCRER 1453
QY 635 KEKIYSFEKFECKE--KDCDE--NTCKNKSEYKMWIDLKSEYKQVDKYTKDKNK 689
Db 1454 KKK---EEVERDKDEYEGCEKEKNGKCVTACKAYKEYITNKKEEYDSQKGFDEKTE 1510
QY 690 KMYNDIDEVKNKEANYLKEKSKCKVDNPFDDKIFNESPEYEDMCKK--CDSEIKYLNFI 747
Db 1511 KK-QGYEDYSEKQASEYLKEK-----CITKSSCNMKWKVTRI 1545
QY 748 -KYPKTKHDIYDIDTSDTFG-----DGTPISTANINQOQCKDTSNTCSNSETSDSPVS 801
Db 1546 SNYTWNPHTYD-----TENLGIKCECPSPFCTIVDGLISFQ-----NSSYANGCKWKYG 1596
QY 802 HEPESDAAINVEKLSGDESSSETRGILDINDPSVTNNVNEVHDASNTQGSVNSDTITNG 861
Db 1597 KMSQGTEDWDCSKKSGEGNEDGVVVCIPRRRRLYVKNLQDLTGESLV----DLRKA 1652
QY 862 HBESSLNRTNA-ODIKIGRSGNEQSDNOENSHSS-----DMSGSLTIGQVPSEDNTQNT 916
Db 1653 FTKCAAIETFFAWHEFKEREKEKNEQDVQYKSVLENLOKLNKGIDDEFKRQMF 1712
QY 917 YDSQNPHRDTPNALASLPSSDDKINEIEGFDSSRSDSENGRGDTTNTDHDVTRTIVISERRV 976

Db 1713 YTFAD-YRDI-----CLGKDIGNVDGINEKIDTILQKNGKPNIBEYKWK-----WQK 1760

Qy 977 NSHDPFIRNGMANNNAHXYTOIRNGGIIRGOESAGNSVNYK-----DNPRESNF 1027

Db 1761 HGHE-IWEGMLCALSNTYETKEMDKELRNKLTQKNGKNTYDVTISGGPIGNTKLEK 1819

Qy 1028 SSENDHKNTOQVNSRDTKVRBEEIIKLSKONKNNEYSMEYCTYSDERNSSPGPC--- 1083

Db 1820 ASRPPPPFRWLEEWADEFCKRTRTHKLEKI--QNECKGVSGTNQC--DDGPGDCDEMCPKD 1875

Qy 1084 -SREERKLCG-----QISDYCLKYFNFSIYEYNCIK 1115

Db 1876 GSFTFKLCSAKSCRYKWKISRKKEEFQKQKYYNEIDDVKRNSDNIYGDFLFTLD 1935

Qy 1116 SEIKSPYKCFKSGQSSIPYFAAGGILVIVILLSSASRMGKSNSEYDICESNIEATF- 1174

Db 1936 QQYKSVLEFKVKGPCSI-----NNNECKIDFNKPKDTFG 1973

Qy 1175 -----BEN-----NYLNKLSRIFNOEVOETN-----ISDYSEYN 1203

Db 1974 HAKNCGPCSBIRFKCIEDNSNWVTNTCNKTFKFTEDNKDTKEDSBLGLMLISDNTVQN 2033

Qy 1204 Y 1204

Db 2034 F 2034

RESULT 36

AAW22475

ID AAW22475 standard; protein; 3060 AA.

XX AC AAW22475;

DT 12-SEP-1997 (first entry)

XX Plasmodium var-7.

XX DBL gene family; SABP; stalic acid binding protein; vaccine; therapy;

KW Duffy binding like gene; Duffy antigen binding protein; erythrocyte;

KW DABP; merozoite; malaria; var-1; var-2; var-3; var-7; immune response;

XX Plasmodium.

XX Plasmodium vivax.

OS Plasmodium falciparum.

XX PN W09640766-A2.

XX 19-DEC-1996.

XX 07-JUN-1996; 96WO-US009508.

XX 07-JUN-1995; 95US-00487826.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Sim KL, Chitnis C, Miller LH, Peterson DS, Su X, Welles TE;

XX WPI; 1997-052231/05.

XX N-PSDB; AAT72882.

XX New malaria vaccines - contains cysteine-rich DBL family protein binding

PT domains homologous domains of the Duffy and sialic acid binding proteins.

XX Claim 8; Page 61-67; 96pp; English.

XX This sequence represents var-7 of Plasmodium. Var-7 belongs to the Duffy

CC binding like (DBL) family of genes which have homology to the Duffy

CC antigen binding protein (DABP) and sialic acid binding protein (SABP)

CC conserved regions (see AAT72889 and AAT72888 respectively). The var

CC family of genes modulate cytoadherence and antigenic variation of

CC Plasmodium infected erythrocytes. SABP and the Duffy antigen binding

CC protein (DABP) are soluble proteins that appear in the culture

CC supernatant after infected erythrocytes release merozoites. DABP and SABP

CC mediate the binding of merozoites and schizonts to the erythrocyte

CC surface. These proteins are necessary for erythrocyte invasion by the

CC parasite. This sequence can be used in the compositions of the invention.

CC The compositions are for the treatment and prevention of malaria, and

CC comprise either a nucleotide sequence or encoded polypeptide of the var-

CC 1, var-2, var-3 or var-7 genes of the DBL gene family, a family of genes

CC having homology with conserved regions of DABP and SABP. The compositions

CC are used for the treatment and prevention of malaria. They are also used

CC in the preparation of vaccines for inducing a protective immune response

CC in a mammal to Plasmodium merozoites (especially Plasmodium falciparum or

CC Plasmodium vivax)

XX

SQ Sequence 3060 AA;

Query Match 7.7%; Score 498.5; DB 2; Length 3060;

Best Local Similarity 18.9%; Pred. No. 1.7e-21;

Matches 317; Conservative 180; Mismatches 520; Indels 661; Gaps 67;

Qy 90 NTKTSYDIIIPPSYSYRNDKFNLSNEDNSGNTNSNNFANTSE-ISIGKDNKQYTFIQKR 148

Db 1309 NKEYTFKQPPPEVATACDCINRSQTEPKKKEENVESACKIVEKILEGKNG-----R 1360

Qy 149 THLPACGIGKRSIKWICRENS--KITVCVDPDRKIQLCVANFLN-----SRLETWEKEKEI 202

Db 1361 TTVGECNPKESYPDWCKNNIDISHDGACMPRRQRKLCYYIAHESQTENIKTDNLDKDA 1420

Qy 203 FLISV-----NTRAKLYNNEGKDRPSIFCNELRNSFSDFRSSFIGDDMD 247

Db 1421 FIKTAAETPLSQWYKSKNDSEAKIL---DRGLIPSOFLRSMYTFGDR-----D 1469

Qy 248 FGGTDRVKGYINTKFSDDYKERNVEKLNLI-----KKEWERNKANLWNHM 294

Db 1470 ICLNTD-----ISKQNDVAKAK--DKIGKFPKSGSKSPSLSRQSEWMTNGEPIWKGM 1522

Qy 295 IV-----NHKGNISKECALIPAEPEPINLWIKENENFL 328

Db 1523 LCALTKYVTDTDNKRKIKNDYSYDKVNQSQGNPSLEEF---AAKQFLRMWIEGSEFC 1579

Qy 329 MEKKRLFLNKKDKCVENKKYEAFCGG---CRLPCSSYTSFMKSKTKQMEVLTN----- 378

Db 1580 AERQKKEIIKDACNEINSTQCCNDAKHRCNQACRACRYQYVENKKKESFGGNTNFFVLKAN 1639

Qy 379 -----LYKKKNSGVDKNNFLNDFLKNKNNDLD-----DFFKNEKEYD 416

Db 1640 VQPDPEYKGYEYKGVQPIQNGEYLLQKDNKKSCMDGNVLSVSPKPEFGYAHKYP 1699

Qy 417 DLCDCRYTATIISFLNGPAKND-----VDIASQINV--NDLRGP---GCNYKSN 461

Db 1700 EKDCD-YQGHVPSIPIPPPPVQVQPEAPTVTVDVCSIVKTLFKDTNNFSDACGLKYGT 1758

Qy 462 NEKSNCTGTFTNKPFGT-----CEPRQTLCLGR----- 492

Db 1759 APSSWKICIPSDTKSGAGATTGKSGDSGSGSICIPPRRRRLYVGLQEWATALPOGEGAAPS 1818

Qy 493 -----TYLLHRGHEEDYKEHLGLGASIVYEAQLLYKYEKEDNLCASI 534

Db 1819 HSRADDLRNAPIQSAAIETFFLDWRYKEEKPKQDGSQQALSQLT--TYSDDEEDPDKL 1877

Qy 535 IQN-----SYADLADIIGKSDIIKDYGKMBEENLNKVNKDKKRNESILKIF 581

Db 1878 LQNGKITPPDFLRLMFTLGDYRDLVHGNTSD--SGNTGNSNNNIVLEASGNKEDMQKI 1936

Qy 582 RE-----KWDENKENYKWSAVL-----KNEKTC----- 607

Db 1937 QEKIEQILPKNGGTPLVPKSSAQTDPKWNHEAESIWKGIMICALTYTEKNPDTSARGDEN 1996

Qy 608 ---KD---YDKF-----QKI----- 616

Db 1997 KIEKDEVEYKFGSTADKHGTASTPTGTQYDYEKKVLEDTSGAKTFSASDTPLS 2056

Qy 617 -----PQFLRWPKWGDGFCERKKEKIYSPESPKVECKK----- 651

Db 2057 DFVLRPPYFRYLEWGNFCKKHKLU---AQIKHECKVEENGSGRRGGITRQYSGDGE 2113

```
QY 652 -----DCDENTCKNCKSEYKKWIDLKKESEYKQYDKYTKDKNKKMYONIDE 697
Db 2114 ACNEMLPKNDGTVPDLLEKPSCAKPSYRKWIESKGEKFEKQKAYEQQDKCVNGS--- 2170
QY 698 VRNKEANVYLK-----EKSK-----CKDVNF-----DDKIFNESPN-----EYED 733
Db 2171 --NKHDNGPCELTLTSSKAKDLTKLGPCKPNNVEGKTIIDDDKTFKHTKDCDPLKFSV 2228
QY 734 MCKK--CDEIKYLNKPKYTKKHDYDI-----DTFSDTFDGTPTISINA 776
Db 2229 NCKKDECDNSKGTD--CRNKNSIDATDIENGVDSTVLEMRVSADSKSGFGDGLGNACRG 2286
QY 777 -----NINEQSQSKO----- 786
Db 2287 AGFEGIRKDEWKCRNVGVVCKPENVNGEAGKHIIQIRALVLRKRWVEYFFDYNNKIKH 2346
QY 787 --TSNTGNSETS-----DSPVSHPEPESDAALNVEKLSGDESSSETRGILD--- 829
Db 2347 KISHRIKNGEISPCIKNCVKWQDKRKWKKEITERFKQYKNDNSDDDNVRSFLETLIP 2406
QY 830 -INDPSVTVNNVNEVHDASNTQG--SVSNTSDITNGHSESSLNRTTNAQDIKIGRSGNEQSD 887
Db 2407 QITDANAKNKVILKSKFGNSCGCSASANEQNKNGEYKDAIDCMLKKLKDKGEC--EKXH 2464
QY 888 NOENSSHSDNSGLTI-----GQVPSD 911
Db 2465 HQTSDECSDTPOQTLEDETLDDETEBAKNMMPKICENVLTKAQOQDEGCVPAPEN 2524
QY 912 N-----TONTYDSQNP 923
Db 2525 SEEPAATDSKETPEQTVLKPREEAVPEPPPPQEKAPAPIPQPPPTPTQLLDNPH 2584
QY 924 -----RDPNLAASL-----PSDD-----KINE----- 941
Db 2585 VLTALVTSTLAWSGIGFATFTFYLLKTKSSVGNLFQILQIPKSDYDPTKLSPNRYI 2644
QY 942 -----IEGFDSSRDS--ENGRGDTSTNTHVRRTNIVSERVNSHDFIRNG 985
Db 2645 PYTSKGYRGKRYIYLEG--DSGTDGTYDHYSDITSSSEYEEMDI-----NDIYVPG 2695
QY 986 MANNAAHQYITOIE-----NNGIIRGOESA-GN--SVNYKDNPKRNSFSEND--- 1032
Db 2696 SP-----KYTLIEVLEPSGNTTASGNNTTASGNNTTASGNNTSPTQNDIQNDGIP 2749
QY 1033 -HKNTIQEYNSRDTKRVREBIIKLSKONKNBYSMYCYTSDERNSSPGCSRE-ERK 1090
Db 2750 SSKITDNEWN-----QLKDSFISQYLQSEPNTEPNM--LGYNVDNNTHTPTSHHVEEK 2802
QY 1091 LCCQISDYCLKYFNFYISYIYNNCKSEIKSPYKCFKSEGOSSIPYPAAGILVIVLL 1150
Db 2803 FIMSIIHDR-----NLFSGEYN-----YDMFNS--GNPFI----- 2830
QY 1151 SSASRMGKSNEEDIGESNTEATFEENNYLNKLSRIFNQVQETNISDYSEYNNK 1208
Db 2831 ----NISDSTNSMDSLTSNNHSPNDKNDLYSGIDLINDALSNGHIDIYDMLKRKEN 2884

RESULT 37
AAV77905
ID AAV77905 standard; protein; 3060 AA.
XX
AC AAV77905;
XX
DT 13-JUN-2000 (first entry)
XX
DE Plasmodium var-7 polypeptide.
XX
KW DBL gene; Duffy-binding like gene; ebl-1; Duffy Antigen Binding Protein;
KW DABP; Sialic Acid Binding Protein; SABB; malaria; vaccine; immunisation;
KW protozoacide; var-7.
XX
OS Plasmodium sp.
```

```
XX
PN US5993827-A.
XX
PD 30-NOV-1999.
XX
PF 07-JUN-1995; 95US-00487826.
XX
PR 10-SEP-1993; 93US-00119677.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Sim KL, Chitnis C, Peterson DS, Su X, Welles TE, Miller LH;
XX
DR WPI; 2000-194198/17.
DR N-PSDB; AA298287.
XX
PT Isolated protein binding domains from Plasmodium vivax and Plasmodium
PT falciparum erythrocyte binding proteins useful for vaccinating against
PT malaria.
XX
PS Disclosure; Col 109-124; 93pp; English.
XX
CC The invention relates to ebl-1 polypeptides that are encoded by the DBL
CC (Duffy-binding like) gene family. The ebl-1 proteins are substantially
CC identical to the Duffy Antigen Binding Protein (DABP) and Sialic Acid
CC Binding Protein (SABP), which are soluble proteins that appear in the
CC culture supernatant after erythrocytes infected with malaria release
CC merozoites. Immunohistochemical studies indicate that DABP and SABP are the
CC respective ligands for Plasmodium vivax and Plasmodium falciparum Duffy
CC and sialic acid receptors on erythrocytes. The ebl-1 polypeptides may be
CC used to vaccinate against malaria, especially caused by P. falciparum.
CC Immunization with the polypeptide provides effective protection against
CC malaria. The present sequence represents the var-7 polypeptide
XX
SQ Sequence 3060 AA;
XX
Query Match 7.7%; Score 498.5; DB 3; Length 3060;
Best Local Similarity 18.9%; Pred. No. 1.7e-21;
Matches 317; Conservative 180; Mismatches 520; Indels 661; Gaps 67;
QY 90 NKTYSYDIIPPSYVRNDKFNLSNEDNSGNTNSNFANTSE-IGIKDNKQVTFIQKR 148
Db 1309 NKEYTFKQPPPEYATACDINRSQTEPKKEENVESACKIVKILEGKNG-----R 1360
QY 149 THLFACGIRKSKIKWICRENS--KITVCVPDKIQLCVANFLN----SRLETWEKPEI 202
Db 1361 TTVGECNPKESYPMDCKNIDISHDGCWPPRRQKLCLYYIAHESQTENIKTDNLDKA 1420
QY 203 FLISV-----NTEAKLLYNKNEGKDPISFCNELRNSFSPRSSFIGDDMD 247
Db 1421 FIKTAAAEFTLSQWYYSKNDSEAKIL---DRGLIPSQFLRSMMYTFGDYR-----D 1469
QY 248 FGQNTDRVKGTYNTKPSDYKKEKVEKLNMI-----KKEWKEKNKANLWNHM 294
Db 1470 ICLNTD-----ISKQNDVAKAK--DKIGKFSKOGSKSPSGLSRQEWKTNQPEIKWGM 1522
QY 295 IV-----NHKGNISKECALIPAEFPQINLWIKWENENFL 328
Db 1523 LCALTXYVTTDNRKIKNDYSYDKVYQNGNPSLEEF---AAKQFLRWMIENGGEFC 1579
QY 329 MEKKRLFLNKKCKVENKKEACFGG---CRLPCSSYTSFMKSKTKQMEVLTN----- 378
Db 1580 AERQKKEIKKDACNEINSTQCNDAKRCNQACRAYQYEVYENKKEKFSQTNFVLKAN 1639
QY 379 -----LYKKKNSGVDKNNFLNDFKQKNKNDL-----DFFKNEKEYD 416
Db 1640 VQPDPEYKGYEKDGVQPIQGNVEYLLQKCDNNKSCMDGNVLSVPKPEKPFQYAHKYP 1699
QY 417 DLCDRCYATATIKSFLNGPAKND-----VDIASQINLV--NDLRGF----GCNYSKN 461
Db 1700 EKDCD-YQGKHPVSIPIPPPPVQPEAPTVTVDCVSIKTLFKDTNNTNFSDACGLKYGKT 1758
QY 462 NEKSNWCTGTFTNKFPGT-----CEPFRQTLCLGR----- 492
```


1067 EKHDGKDNPKYTGDESKPAYK--KL RADWEANRHQVWRAMKCATKGI--CPGMPV 1121
312 EE--PQINLWIKENENFLMEKRLFLNIIKDKCUE--NKYIEAC-----FGGRLPCSS 361
1122 DDYIPQRLRWMTWAWEYKCAQOEYDKLKKI CADCMKSGDKGKCTOGVDGCRKAACDK 1181
362 YTGFMKSKTKQMEVLNLY--KKNSGVDXN-----NFLNDLPK-- 398
1182 YKEIEKWNQWKISDKYNLLYLQATTSNTNPGRTVLGDDDDPYOQMVDFLTPHKASI 1241
399 ----- 398
1242 AARVLVKRAAGSPTEIAAAAPITPYSTAAGYIHQEIYGGCQEQGTQCEKKHGATSTTT 1301
399 KNNKNDLDDFFKNEKEYDDLCDRYTATIIKSLFNGPAKNDVDIASQINVD--LRFG- 455
1302 KENKE--YTQKPPPEYATACD-----INRSOTEPKKEENVESACKIVEKILEGKNG 1354
456 -----CNYKSNNEKSNCTGTFTNKPFGTCEPPRQTLCLGRTYLLHRGHEEDYK--EH 507
1355 RTTVGECNPK-ESYPDWDCNKNIDISHGACMPRRQKJCL--YYIAHESQTENIKTDNN 1411
508 LLGASIEYA---QLLYKYKEKDNALCSIION-----SYADLADIIRGSDI 551
1412 LKDAFIKTAATAETFLSQYKSKNDSEAKILDRGLIPSQFLRSMWYTFGDRDILCLNTDI 1471
552 IKDYIG-KQMEENLNK-VNKDKRNEESLKIPEKWDENKENVNVKMSAVL-----K 602
1472 SKQNDVAKAKDKIGKGFSGKSGSPSGLS--RQEWKTKNGPEIWKGMCLCATKYVTDTD 1529
603 NKETCKD---YDK-----PQKIPQFLRPFKEWDDDFCEKREKIKYISFESFKV 646
1530 NKKIKNDYSDKVNQSONGNPSLEFPAAPQFLRMWIEGEEFCAERQKK--ENIIKD 1586
647 ECKK---KDCDE--NTCKNCKSEYKKWIDLKSEYEKQVDKYT-----KDKNKKMYD 693
1587 ACNEINSTQCNDKAHRCNACRAYQEYVENKKKKEFGSGTNNFVLKANVQPPQPEYKGYE 1646
694 NIDEVKNKEANVILKEKSECKDVNFDKIFNPSNE-----YEDMKCKDEIKYL 744
1647 YKGVQPIQNEYLLOKQDNCKNCKSCMDGNVLSVPKEKPGKYAHKYPEK-DCYQGHV 1705
745 NETKYPK-----TKHDIYDIDTFSDT-----FGDGTPIISANINEQ 781
1706 PSIPPPPPVQPOPEARTVTVDCSIVKTLFKDTNNFSDACGLKYKGTAPSSWKCPISPT 1765
782 QSGKDTNTGNTSETSDSPV-----SHEPESD----- 807
1766 KSGAG-ATTCKSGSDSGSICIPRRRLYYVYKLEWATALPQGEGAAPSHSRADDLRNF 1824
808 ----AAINV-----EKLSDGESSETRGIL-----DINDPSVTNNVNEVHDASNTQ 849
1825 IQSAAIETFLWDRYKEEKPOGDSQOALSQLTSTYSDDEEDPP-----DKLLQN 1875
850 GSV-----SNTSDI--TNGHSESSLNRTTNAQDIKIGRSGNEQ 885
1876 GKIPPPFLRMFTLGDYRDLVHGNTSOGTNG-----SNNNNVILEASGNKE 1926
886 SDN--QENSSHSDNSGSLTI-----QQVPE--DNTQN-----TYDSQNPHRDT 926
1927 DMQKIQEKIEQILPKNGGTFLPVKSSAQTDPKWNEHAESIKWGMICALTYTEKNPDTSA 1986
927 PNALASLPDDKINEIGFPOSSRDSSENGRGDTTNTSHDVRRTNIVSERVNSHDFTRNGM 986
1987 RGDENKIEKDEEV--YEKFFGSGTADKHGTASTTGT----- 2020
987 ANNAHLOYITQIENNGIIRGOEESAGNSVNYKDNPKRNSFSSNDHKNIQEVNSRDTK 1046
2021 -----YKQDYKVKLEDTSGAKTPSASDTPLLSDFLVRPPYFRYLEWGNQFCK 2072
1047 RVREBIKLSKQNK-----CNNEYSMEYCTYSD-ERNSSPGP 1082
2073 KRKHQAQIKHECKVEBNGGSGRRGGITROYSGDGEACNEMLPKNDGTVPDLEKPSCAKP 2132

QY 1083 CSR-----BERKKLCCQISDYCLKYFNFYFYSIEYNYCIKSEIKSPY 1123
Db 2133 CSSYRKWIESKGKEFEKQKAYEQQDKCVNGSNKHDNGFCETLTITSSKADF 2185

RESULT 39

AAW22482
ID AAW22482 standard; protein; 2710 AA.
XX AC AAW22482;
XX 07-OCT-1997 (first entry)
XX Plasmodium Proj3.
XX DBL gene family; SABB; sialic acid binding protein; vaccine; therapy;
KW Duffy binding like gene; Duffy antigen binding protein; erythrocyte;
KW DABP; merozoite; malaria; var-1; var-2; var-3; var-7; immune response;
XX Plasmodium.
XX Plasmodium falciparum.
XX OS WO9640766-A2.
XX PN 19-DEC-1996.
XX PD 07-JUN-1996; 96WO-US009508.
XX PF 07-JUN-1995; 95US-00487826.
XX PR (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PA Sim KL, Chitnis C, Miller LH, Peterson DS, Su X, Wellems TE;
XX WPI: 1997-052231/05.
XX N-PSDB; AAT72897.
XX New malaria vaccines - contains cysteine-rich DBL family protein binding domains homologous domains of the Duffy and sialic acid binding proteins.
XX Disclosure; Page 50-56; 96pp; English.

XX This sequence represents Proj3 of Plasmodium. Proj3 belongs to the Duffy binding like (DBL) family of genes which have homology to the Duffy antigen binding protein (DABP) and sialic acid binding protein (SABB) conserved regions (see AAT72889 and AAT72888 respectively). The var family of genes modulate cytoadherence and antigenic variation of Plasmodium infected erythrocytes. SABB and the Duffy antigen binding protein (DABP) are soluble proteins that appear in the culture supernatant after infected erythrocytes release merozoites. DABP and SABB mediate the binding of merozoites and schizonts to the erythrocyte surface. These proteins are necessary for erythrocyte invasion by the parasite. This sequence can be used in the compositions of the invention. The compositions are for the treatment and prevention of malaria, and comprise either a nucleotide sequence or encoded polypeptide of the var-1, var-2, var-3 or var-7 genes of the DBL gene family, a family of genes having homology with conserved regions of DABP and SABB. The compositions are used for the treatment and prevention of malaria. They are also used in the preparation of vaccines for inducing a protective immune response in a mammal to Plasmodium merozoites (especially Plasmodium falciparum or Plasmodium vivax)
XX Sequence 2710 AA;
SQ

Query Match 7.6%; Score 495.5; DB 2; Length 2710;
Best Local Similarity 19.5%; Pred. No. 2.2e-21;
Matches 268; Conservative 170; Mismatches 482; Indels 453; Gaps 57;
QY 112 LSENNEDNGNTNPNFANTSEISIGNDKQYTFIQKRTLHACGIR-KRSIKWICRENSE 170
Db 912 INENYSDNRSGNSGGPCT-----GKGDH-----GGVRMRIGTWSNIEGKK 953

171 KIT---VCVPRDKIQLCVANFLNSRLTM---EKFKIFLISVNTAKL----- 213
Db QTSYKVNFLPRRHHMCTSNLENLDVGSVTKNDKASHLLGDVQLAAKTDAEIIKRYKD 1013
214 ---LYNKGEGDPSIFCHELNSPDSFGDDM--DFGNTDRVKGII----- 259
Db QNNIQLTDPLOQDQEAACRAVRYSFADLGDIIIGRDMWDEKSDTDMETRLITVFNKIK 1073
260 ---NTKFSYDYKQVKEKLNKIKGWEKKNKANLNMHMIIVNKHGNIKECALIPA 311
Db EKHDGIDKNPKYTGDESKPAYK--KLADWWEANRHQVWRMKCATKGII---CPGMPV 1128
312 BE--PQNLWIKENENFLMEKGLFLNIDKQVVE--NKYFAC-----FGGRLPCSS 361
Db DDYIPQRLRWTEWAERYCAQSQEYDKLKICADCMKSGDKCTQGDVDCGKCAACDK 1188
362 YTSFMKSKTQMEVLTNLY-----KKNSGVGDKN-----NFLDLFK--- 398
Db YKEIEKWEQWRKISDKYNLLYLQAKTITNPGRTVLGDDDPDYQQWVDFLTPHKASI 1248
399 ----- 398
1249 AARVLVKAAGSPTEIAAAPTPTAAGYIHOEIGVGGCQEQTFCEKHGATSTTT 1308
399 KNNKNDLDDFFKNEKYDDLDCCRYTATIIKSFNGPAKNDVDIASQINVD--LRGFG- 455
Db KENKE--YTFKQPPPEYATACDC-----INRSQTEEPKKEENVESACKIVEKILEGKNG 1361
456 -----CNYKSNKSNKCTGTFNKPFGTCPEPRRQTLCLGRYLLHRGHEEDYK--EH 507
Db RTTVGECNPK--ESYPMDCWCKNIDISHGACMPPRRQKLC--YYIAHESQTENIKTDN 1418
508 LLAGASIYEA---QLLKYYKEKEDENALCSIIQN-----SVADLADIIGSDI 551
Db LKDAFIKTAARETFLSQYKSKNDSEAKILDRGLIPSFRLSMYTFGDRYDLCITDI 1478
552 IKDYIG--KMEENLNK--VNKDKRNEESLKIPEKRWNDENKENVYKMSAVL-----K 602
Db SKQNDVAKAKDKIGKFFSKDGSKSPGLS--RQEMWKTNGPELWKGMALCALTYVTD 1536
603 NKETCKD---YDK-----FQKJPOFLWFKWGGDDFCERKKEIYFESFKV 646
Db NKRIKNDYDYKVNQSONGNPSLEEFAPKQFLRWMIWGEFCAEROKK--ENIKD 1593
647 ECKK---KDCDE--NTCKNKCSEYKWKIDLKSEYKQVDKYT-----KDKNKKMYD 693
Db ACNEINSTQCNDAKHCNOACRAYQYVENKKEFGSQTNFVLKANVQDPPEYKGYE 1653
694 NIDVKNKEANVYLKEKSECKDVNFDKIFNESPNB-----YEDCKKDEIKYL 744
Db YKGVQPIQGVNLLQKCDNNKSCMDGNVLSVSPKEKPGYAHKYPEKC--DCYQKGHV 1712
745 NEIKYPK-----TKHIYDIDTFSDT-----FGDGTPLISINANINEQ 781
Db PSIPPPPPVQPOEAPTVTVDVGSIVKTLFKDTNFSADGLKYGKTAPSSWKCPISDT 1772
782 QSGKDTNKGTSFSDSPV-----SHEPESD----- 807
Db KSGAG--ATTGSGSDSGSICIPRRRLRYGVKQLQEWATALPQGGGAAPSHRADDLNRAF 1831
808 ---AAINV-----EKLSDGESSETRGII-----DINDPSVTNNVNEHVASNTQ 849
Db IQSAAITFLWDRYKEKFPQDGSQALSQLTSTSYDDEEDPP-----DKLLQN 1882
850 GSV-----SNTSDI--TNGHSESSLNRTTNAQDIKIGRSGNEQ 885
Db GKIPDPFLRLMFLTYLGDYRDLVHGNTSDSGNTG-----SNNNIVLEASGNKE 1933
886 SDN--QENSSHSSDNGSLTI-----QGVSE---DNTQN-----TYDSQNPDRDT 926
Db DMOKIQEKIEQIPLKNGGTFLVPKSSAQTPDKWNEHAESIWKGMICALTYTEKNPDTSA 1993
927 PNALASLPDDKINEIEGFDSSRDSNGRGDTTNTHTDVRTNIVSRRVNSHDFIRNGM 986

Db 1994 RGDENKLEKDEV--YKFFGSGTADKHGTASTPTGT----- 2027
Qy 987 ANNAHQAQYITQJENNGIIRGOESAGNSVNYKNDPKRSPSSSDHKKHIOEYNSRDTK 1046
Db 2028 -----YKQVDYKVKLEDTSGAKTPSASDTPLLSDFLVPPYFRYLEEGWQNFCK 2079
Qy 1047 RVREIITKLKQNK-----CNNEYSMEYCTYSD-ERNSSPGP 1082
Db 2080 KRKHQAQIKHECKVEBGGSGRRGGITROYSGDGEACNEMLPKNDGTVPLEKPSCAKP 2139
Qy 1083 CSR-----BERKLCQOISDYCLKYFNFYFIEYNYCNKISIKSPEY 1123
Db 2140 CSSYRWKIESKGEFEKQKAYEQKQKCVNGSKNGKNDGFCETLTTSSKAKDF 2192

RESULT 40
AAY77904
ID AAY77904 standard; protein; 2710 AA.
XX AAY77904;
XX AC
XX 13-JUN-2000 (first entry)
DT P. falciparum Proj3 binding domain polypeptide.
DE DBL gene; Duffy-binding like gene; ebl-1; Duffy Antigen Binding Protein;
KW DABP; Sialic Acid Binding Protein; SAPP; malaria; vaccine; immunisation;
KW protozoacide; Proj3.
XX Plasmodium falciparum.
XX US5993827-A.
XX 30-NOV-1999.
XX 07-JUN-1995; 95US-00487826.
XX 10-SEP-1993; 93US-00119677.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Sim Kl, Chitnis C, Peterson DS, Su X, Wellem TE, Miller LH;
XX WPI; 2000-194198/17.
XX N-PSDB; AAZ98286.
PT Isolated protein binding domains from Plasmodium vivax and Plasmodium
PT falciparum erythrocyte binding proteins useful for vaccinating against
PT malaria.
XX Disclosure; Col 79-92; 93pp; English.
XX The invention relates to ebl-1 polypeptides that are encoded by the DBL
XX (Duffy-binding like) gene family. The ebl-1 proteins are substantially
XX identical to the Duffy Antigen Binding Protein (DABP) and Sialic Acid
XX Binding Protein (SAPP), which are soluble proteins that appear in the
XX culture supernatant after erythrocytes infected with malaria release
XX merozoites. Immunochemical studies indicate that DABP and SAPP are the
XX respective ligands for Plasmodium vivax and Plasmodium falciparum Duffy
XX and sialic acid receptors on erythrocytes. The ebl-1 polypeptides may be
XX used to vaccinate against malaria, especially caused by P. falciparum.
XX Immunization with the polypeptide provides effective protection against
XX malaria. The present sequence represents the Proj3 binding domain
XX polypeptide
SQ Sequence 2710 AA;

Query Match 7.6%; Score 495.5; DB 3; Length 2710;
Best Local Similarity 19.3%; Pred. No. 2.2e-21;
Matches 268; Conservative 170; Mismatches 482; Indels 453; Gaps 57;
Qy 112 LSENEDSNGTNGNNTFANTSEISIGKDNKQYTFQKRTHLFACGK-RKSIKWICRENS 170

```
Db 912 INENYSNDRSGNGPCT-----GKGDH-----GVRRRIGTEWSNEGKK 953
QY 171 KIT---VCVPRDKIQLCVANFLNSRLTM---BKFEIIFLISVNTBAKL-----213
Db 954 QTSYKVNFLPPREHMCTSNLENLDVGSVTKNDKASHLLGDVQLAAKTDAAEIIKRYKD 1013
QY 214 -----LYNKNEGKOPSIFCNELRNSPSDFRSSIFGDDM---DFGNTDRVKGYI-----259
Db 1014 QNNIQLTDPITQQKQDOEAMCAVRYSFADLGGIIRGRDMDWDEKSDTDMETRLTIVFKNIK 1073
QY 260 -----NTKFSDDYKKEKNVLEKLNIIKKEWKEKNKANLWNHIVNHKGISKECALIPA 311
Db 1074 EKHDGIGKDNPKYTGDESKPAYK--KLURADWEANRHQVWRAMKCATKGII---CPGMPV 1128
QY 312 EE--PQINLMIKEWENFLMEKKRLFLNIKDKCQE--NKKYEAC-----FGCRLPCSS 361
Db 1129 DDYIPQRLRWMTWEAWYCKAQOQEYDKLKKI CADCMSKDGKCTQGDVDCGKCAACDK 1188
QY 362 YTSFMKSKTQMEVLTNLY-----KKKNSGVDKN-----NFLNDLPK---398
Db 1189 YKEIEKWNQWRKISKDKNLLYLQAKTTSTNPRTVLGDDDPDYQQWVDFLPIHKASI 1248
QY 399 -----398
Db 1249 AARVLKRAAGSPTEIAAAAPITPYSTAAGYIHQEIYGCGCQBTQFCEKKHGATSTTT 1308
QY 399 KNNKNDLDDFFKNEKEYDDLCDRYTATIIKSLFNGPAKNDVDIASQINVD--LRGFG- 455
Db 1309 KENKE--YTFKQPPPEVATADC-----INRSQTEBPKEEENVEACKIVEKILEGKNG 1361
QY 456 -----CNYKSNNEKSNWCTGTFTNKPFGTCEPPRQTLCLGRYTLHLRHGHEBYK--EH 507
Db 1362 RTTVGECNPK--ESYPWDCKNNIDISHGACMPPRQKCL--YYIAHESQTENIKTDN 1418
QY 508 LLAGASIVEA---QLLKYYKEDENALCSIIQN-----SYADLADIIGSDI 551
Db 1419 LKDAFRTAAAEFTLSWQYYKSKNDSEAKILDRGLPSQFLRSMYTFGDIYRDICLNTDI 1478
QY 552 IKDYYG-KKMEENLNK-VNKDKRNEBSLKFREKWDENKENVWKMVASL-----K 602
Db 1479 SKQNDVAKAKDKIGKFFSKDGSKPSGLS--RQEWKNTNGPEIWKMGLCALTKYVTDTD 1536
QY 603 NKETCKD---YDK-----FQKIPQFLRMFKWGGDDFCEKRKEKIYSPESFKV 646
Db 1537 NKRKIKNDYSYDKVNSQNGNPSLEFAAKPQFLRWMIENGEBFAERQKK---ENIIO 1593
QY 647 ECKK----KDCDE--NTCKNKSEYKWKIDLKSEYEKQVDKYT-----KDKNKKMYD 693
Db 1594 ACNEINSTQCNDAKHRCNOACRAYOBYENVENKKFSGQTNFVLKANVQOPDPEYKGYE 1653
QY 694 NIDEVKNKEANVYLKEKSECKDVNFDKIFNESPNE-----YEDMCKKDEIKYL 744
Db 1654 YKDGVPQIQGNEVLLQKCDNNKCSMDGNVLSVPKEKPGKYAHKYPEKC--DCYQGHV 1712
QY 745 NEIKYPK-----TKHDIYDIDTFSDT-----FGDGTPIISANINEQ 781
Db 1713 PSIPPPPPVQOPQPEAPTIVTVDCSVIKVTLFKDTNNPFSDACGLKYGTAPSSWKIPSDT 1772
QY 782 QSGKOTSNIGNSETSDFV-----SHEPESD-----807
Db 1773 KSAG-ATTGKSGSDSGSICIPRRRLYYVKLQEWATALPQEGGAAPSHSRADDLRNF 1831
QY 808 ---AAINV-----EKLSGDESSSETRGIL-----DINDPSVTNNVNEVHDASNTQ 849
Db 1832 IQSAALETFFLWDRYKEEKPGQDGSQOALSQITSYSDDEEDPP-----DKLLQ 1882
QY 850 GSV-----SNTSDI--TNHSESSLNRTTNAQDIKIGRSGNEQ 885
Db 1883 GKIPDPFLRLMYFTLGDYRDLVHGNTSDSGNTNG-----SNNNNIVLEASGNKE 1933
QY 886 SDN--QENSSHSDNSGSLTI-----GOVPSE---DNTQN-----TYDSQNPDRDT 926
```

```
Db 1934 DMQKIQEKIEQILPKNGGTPLVPKSAQTPDKWNEHAESIWKGMICALTYTEKNPDPTSA 1993
QY 927 PNALASLPDDDDKINIEGFDSSRDSSENGRGDTTSTNTHDVRRTNIVSERRVNSHDFIRNGM 986
Db 1994 RGDENKIEKDDEV--YEKFFGSTADKHGTASTPTGT-----2027
QY 987 ANNNAAHQVITQIENNGIIRGOEESAGNSVNYKDNPKRNSFSENHKKNIQIYNSRDTK 1046
Db 2028 -----YKTQDYEKVKLEDTSGAKTPSASDTPLLSDFLRPPPYFRYLEEWGQNFCK 2079
QY 1047 RVREEEIILKSKONK-----CKNEYSMEYCTYSD--ERNSSPGP 1082
Db 2080 KRKHKLAIQIKHECKVVEENGSGRRGIGITROYSGDGEACNEMLPKNDGTVPDLEKPSCAKP 2139
QY 1083 CSR-----BERKKLCCQISDYCLKYFNFYISIEYINCICKSEIKSPEY 1123
Db 2140 CSSYRKWIESKGKEFEKQEKAYEQQKDKCVNGSNKHDNGFCETLTITSSKAKDP 2192
```

Search completed: November 21, 2005, 20:33:45
Job time : 230 secs